

K1

(19)



Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11)

**EP 0 477 972 B1**

(12)

**EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention  
of the grant of the patent:  
04.09.1996 Bulletin 1996/36

(51) Int Cl.<sup>6</sup>: **C12Q 1/70**, **C12Q 1/68**,  
**C12P 19/34**  
// **C07H21/04**

(21) Application number: **91116561.1**

(22) Date of filing: **27.09.1991**

(54) **Nucleotide sequences useful as type-specific probes, PCR primers and LCR probes for the amplification and detection of human papilloma virus, and related kits and methods**

Nukleotid-Sequenzen nützlich als typenspezifische Sonden, PCR Primers und LCR Sonden zur Amplifikation und zum Nachweis von humanem Papillomavirus, sowie dazu verwendete Kits und Verfahren

Séquences nucléotidiques utiles comme sondes spécifiques du type amorces de PCR et sondes pour l'amplification et détection du virus-papilloma humain, et kits et procédés utilisés dans ce but

(84) Designated Contracting States:  
**AT BE CH DE DK ES FR GB GR IT LI NL SE**

(30) Priority: **28.09.1990 US 589948**  
**28.09.1990 US 590105**  
**28.09.1990 US 590253**

(43) Date of publication of application:  
**01.04.1992 Bulletin 1992/14**

(73) Proprietor: **ABBOTT LABORATORIES**  
**Abbott Park, Illinois 60064-3500 (US)**

(72) Inventors:  
• **Joseph, Jeffrey L.**  
**Cherry Hill, New Jersey 08002 (US)**  
• **Bouma, Stanley R.**  
**Mundelein, Illinois 60060 (US)**

• **Marshall, Ronald L.**  
**Zion, Illinois 60099 (US)**  
• **Laffler, Thomas G.**  
**Libertyville, Illinois 60048 (US)**

(74) Representative: **Modiano, Guido, Dr.-Ing. et al**  
**Modiano & Associati S.r.l.**  
**Via Meravigli, 16**  
**20123 Milano (IT)**

(56) References cited:  
**EP-A- 0 301 968** **EP-A- 0 320 308**  
**EP-A- 0 402 132** **EP-A- 0 425 995**  
**WO-A-87/05630** **WO-A-89/05357**  
**WO-A-89/09940** **WO-A-90/02821**  
**WO-A-91/10675**

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

**EP 0 477 972 B1**

## Description

This invention relates generally to human papilloma virus, and more particularly, relates to nucleotide sequences of short strands of human papilloma virus which can be amplified and/or used to determine the presence of human papilloma virus products in a test sample, and some of which also can be amplified and/or used to determine the specific type of human papilloma virus of types 16 and 18 present in the test sample.

Human papilloma virus (HPV) is recognized as a venereally-transmitted disease of the anogenital tract which often is associated with the pathogenesis of cervical cancer and its precursor lesions. More than 56 types of HPV have been characterized. Of these, at least 21 types infect the anogenital tract. L. Gregoire et al., *J. Clin. Micro* 27 (12): 2660-2665 (1989). These mucosotropic viruses are associated most frequently with benign condyloma or latent infections. However, the presence of HPV in premalignant lesions and invasive cancers, particularly of the cervix, may reflect the oncogenic potential of these viruses. See P. M. Howley, in *Important Advances in Oncology*, D. T. DeVita, Jr. et al., eds., J. B. Lippincott, Philadelphia, PA (1987) at pages 55-73.

Certain HPV types, namely, HPV type 16 and type 18, and to a lesser extent HPV types 31, 33 and 35, are found in a high proportion of invasive cervical cancers and their metastases. However, many HPV types which infect the anogenital tract, such as HPV types 6 and 11, are found most commonly in benign condyloma and only rarely are found in invasive cancers. HPV detected in the anogenital tract can be classified broadly as low risk papilloma viruses ( HPV types 6 and 11), intermediate risk papilloma viruses ( HPV types 31, 33 and 35) or high risk papilloma viruses ( HPV types 16 and 18), based on the association of the particular HPV type with malignancy. A. T. Lorincz et al., *J. Nat'l Cancer Inst.* 79:671 (1987). Thus, the detection of the presence of HPV and the determination of the specific type of HPV can provide a diagnostic and prognostic tool useful for determining the clinical significance associated with certain HPV types. The early detection of HPV by sensitive and specific reagents and methodologies also could provide earlier therapeutic management and counseling.

A need therefore exists for accurate and reliable methods to identify and type HPV in clinical specimens. However, known polyclonal antisera prepared by immunizing animals with disrupted virions are capable of detecting HPV antigens in only about 30-70% of cutaneous and mucosal warts. Further, the antisera are broadly cross-reactive. Available immunological tests have two major drawbacks. First, only well-differentiated cells apparently are capable of viral antigen expression. HPV-infected tissues which show higher degrees of neoplasia, such as carcinoma *in situ*, rarely contain HPV antigen. Thus, the further the development of the malignancy, the smaller the amount of detectable virus in the tested tissue. Secondly, these immunological tests are unable to identify specific viral types.

It is known that papilloma viruses share amino acid sequences in the major capsid proteins. See, for example, C. C. Baker, in *The Papovaviridae* (Vol. 2), P. M. Howley and N. P. Salzman, eds., Plenum Publ. Corp., New York (1987) at pages 321-385. The DNAs of this virus cross-hybridize, indicating homologous sequences. M. F. Law et al., *J. Virol.* 58:225-229 (1979). Thus, molecular hybridization techniques have been developed as a more sensitive and specific means of detecting and differentiating HPV DNA and RNA in clinical specimens. See A. T. Lorincz, *Obstetrics and Gynecol. Clinics of N. America* 14:451 (1987).

Sequences specific for the DNA and RNA of human papilloma virus are known and have been published. See, for example, PCT application No. WO 89/69940 published October 19, 1989, PCT application No. WO 86/05816 published October 9, 1986 and European Patent Application No. 0 301 968 published February 1, 1989.

The molecular hybridization techniques used to detect homologous DNA sequences are sensitive and can be highly specific if used with probes which bind to nucleic acid sequences which are unique to a particular HPV type. However, the concentration of total viral DNA in a given clinical sample may be below the limit of sensitivity of the test. For example, the amount of viral DNA in dysplastic cervical lesions is reduced with increasing dysplasia.

To overcome this problem of sensitivity, viral DNA sequences can be amplified by using, for example, the polymerase chain reaction (PCR) or the ligase chain reaction (LCR) techniques. The products thus obtained can be identified by using conventional hybridization techniques for identification of virus types, such as Southern blotting. See C. Oste, *Biotechniques* 6:163 (1988), K. B. Mullis, U. S. Patent No. 4,683,202, and EP-A-320 308 (BioTechnica).

Both PCR and LCR serve to amplify the DNA present in a test sample to detectable levels. In practice, the level of sensitivity is about 50 to 100 copies per sample. The next most sensitive technique is dot-blot, which can detect about 10,000 molecules, while Southern blot reliably detects about 100,000 copies of DNA per sample.

Thus, the appropriate diagnosis of HPV may require two steps. In one strategy, the presence of a clinically relevant type of HPV is first detected with a group-specific primer. After the presence of HPV is detected, differentiation between types can be performed by using a type-specific probe having low homology between the HPVs of the group. Alternatively, differentiation can be performed using a mixture of type-specific probes at the outset, provided these probes amplify DNA independently of each other, and that they can be detected independently. In the past, such tasks were attempted using specific antibodies. In general, nucleic acid probes and primers allow greater discrimination among subtypes than do antibodies. The use of DNA-based tests increases both sensitivity and specificity over prior-art antibody-based tests.

It therefore would be advantageous to provide oligonucleotide strands of DNA which could be amplified and used to detect the presence, if any, of HPV in a test sample. It also would be advantageous to provide short oligonucleotide strands of DNA which could be amplified and used to detect the presence, if any, of specific types of HPV in the test sample. The combined use of oligonucleotide strands would be advantageous for allowing for the specific and sensitive in vitro diagnosis of the presence and specific type of HPV present in test samples.

#### SUMMARY OF THE INVENTION

Oligonucleotides of from about 10 to about 60 nucleotides are provided which can be amplified and used either to detect specific sequences of specific types of human papilloma virus, or consensus regions with high homology among different types. The presence of HPV is determined by contacting the test sample with sequences provided to detect the presence, if any, of HPV types 6, 11, 16, 18, 31, 33 and 61. This may be done with or without prior amplification, for example, by PCR or LCR. Either type-specific or consensus amplification is also possible. Two oligonucleotides are provided if the sequence is to be amplified by PCR, and four oligonucleotides provided if amplification is by LCR, in accordance with these known amplification procedures. After the presence of HPV is detected, the type of HPV present in the sample can be determined by using HPV type-specific probes, by subsequent rounds of PCR, or by LCR. Alternatively, the presence of type-specific HPV can be determined by contacting the test sample directly with type-specific nucleotide sequence provided by the invention for the detection of HPV types 16 and 18. Also provided are methods for using the oligonucleotides and kits for amplifying and detecting the presence of human papilloma virus.

#### BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a photograph of a gel following electrophoresis showing the results when the primers PCR 1 and PCR5 were used to amplify selected plasmids wherein HPV 6 is in lane 1, HPV 11 is in lane 2, HPV 16 is in lane 3, HPV 18 is in lane 4, and HPV 31 is in lane 5, HPV 33 is in lane 6, HPV 61 is in lane 7, and molecular weight standards are in lane 8.

FIG. 2 is a photograph of a gel following electrophoresis showing the results when the primers PCR 1, PCR2, PCR3, PCR4 and PCR5 were used to amplify plasmid p65.16.8 (HPV 16). PCR1 and PCR5 are primers according to the invention.

FIG. 3 is a photograph of the ethidium bromide-stained gels wherein PCR 1 4 and PCR15 are used in conjunction with IWDO to obtain amplified PCR product.

FIG. 4 is a graph of results obtained from performing LCR on  $10^7$  molecules of the selected target using LCR5A, LCR5A', LCR5B and LCR5B'. The rate of reaction of 4-methyl umbelliferone is expressed as fluorescence counts/second/second and plotted against the target HPV type.

FIG. 5 is a graph of results obtained from performing LCR on  $10^7$  molecules of the selected target using LCR6A, LCR6A', LCR6B and LCR6B'. The rate of reaction of 4-methylumbelliferone is expressed as fluorescence counts/second/second and plotted against the target HPV type.

FIG. 6 is a graph of results obtained from performing LCR on  $10^7$  molecules of the selected target using LCR7A, LCR7A', LCR7B and LCR7B'. The rate of reaction of 4-methylumbelliferone is expressed as fluorescence counts/second/second and plotted against the target HPV type.

FIG. 7 is a graph of results obtained from performing LCR on  $10^7$  molecules of the selected target using LCR8A, LCR8A', LCR8B and LCR8B'. The rate of reaction of 4-methylumbelliferone is expressed as fluorescence counts/second/second and plotted against the target HPV type.

#### DETAILED DESCRIPTION OF THE INVENTION

The appropriate diagnosis of HPV requires two sets of conditions. The first enables the detection of all pertinent types, and the second set allows differentiation among them. In the past, such tasks have been attempted using specific antibodies. In general, nucleic acid probes and primers allow greater discrimination among subtypes than do antibodies. Thus, the use of DNA-based tests tends to increase both sensitivity and specificity over antibody-based tests.

U. S. Patents No. 4,683,195 and 4,683,202 teach a method of amplifying DNA sequences by using PCR. This method now is a standard procedure in many molecular biology laboratories. Examples 1-3 which follow below utilize the procedures taught in these two patents and the method as described in the package insert of the commercially-available Gene-Amp™ kit (Document No. 55635-6/89, Perkin-Elmer/Cetus, Emeryville, CA).

In PCR, two complementary polynucleotide strands are amplified by treating the strands with two oligonucleotide primers such that an extension product of each primer is synthesized which is complementary to each nucleic acid strand. The primers are selected such that the extension product of one primer forms a template for the synthesis of an extension product from the other primer once the extension product of the one primer is separated from the template. A chain reaction is maintained by a cycle of denaturing the primer extension products from their templates, treating

the single-stranded molecule generated with the same primers to re-anneal, and allowing the primers to form further extension products. The cycle is repeated for any many times as it takes to increase the target nucleic acid segments to a concentration where they can be detected.

5 The amplified target sequence can be detected by any of several known techniques: for example, by denaturing the double-stranded products formed by PCR, and treating those products with one or more reporter probes which hybridize with the extension products. The reporter probe has a detectable label, and usually is added in excess. The unhybridized reporter probe, therefore, must be separated from the hybridized reporter probe by involving a separation step. In another method of detecting the extension products without reporter probe and a separation step, the extension products are detected by gels stained with ethidium bromide. The diagnosis can be confirmed by transferring the DNA  
10 to nitrocellulose and probing with a probe specific to the HPV type suspected of being present in the sample.

Alternately with PCR, one may take advantage of known restriction sites within the HPV DNA to demonstrate that the amplified DNA contains the expected sequence by examining the cleavage pattern(s) generated with one or more restriction endonucleases. Verifying the authenticity of the amplified sequence may be necessary for two reasons: (1) to ensure that sequences complementary to the amplifying primers are not fortuitously present in cellular DNA which  
15 does not contain HPV DNA, and (2), to identify the type of HPV present in the sample. If the sequences chosen for amplification are conserved among HPV types, then the finding of an amplified product does not implicate a particular HPV type. It also should be possible to predict the size of the amplified product based on the binding positions of the two primers. Thus, when that product is found, one reasonably can be assured that HPV is present. However, two different types of HPV may give the same or different size products. Thus, hybridization should be used to confirm the  
20 identity of the amplified sequence until confidence is built that the interpretation of the results is reliable. It should be pointed out that the PCR technique will identify only closely related, or type-specific sequences in the absence of highly homologous primers, since only a small portion of the genome is analyzed.

Another particularly useful detection technique is described in EP-A-357 011. In this method, a different reporter molecule, e.g. hapten, is attached to each primer. Following amplification, but before denaturation, duplexes can be  
25 detected by "capturing" one hapten (hapten1) with a solid phase coated with anti-hapten1. The separated complex can be detected with a conjugate of label and anti-hapten2, and label associated with the solid phase can be measured.

The Ligase Chain Reaction (LCR) amplifies sections of DNA by copying the section of DNA, and copying the copies of that section of DNA, many times over. This method is described in European Patent Application No. 0 320 308 published June 14, 1989, which is incorporated herein by reference. In this procedure, two probes (for example,  
30 A and B) complementary to immediately adjacent regions of a target sequence are hybridized and ligated. This ligated probe then is denatured away from the target, after which it is hybridized with two additional probes (A' and B') of sense opposite to the initial probes A and B. The secondary probes are themselves then ligated. Subsequent cycles of denaturation/hybridization/ligation create the formation of double-length probes of both sense (+) and antisense (-).

In LCR, the nucleic acid of the sample is provided either as single stranded DNA or as double-stranded DNA which  
35 is denatured to separate the strands. Four probes are utilized: the first two probes (A and B) are the so-called primary probes, and the second two probes (A' and B') are the so-called secondary probes. The first probe (A) is a single strand capable of hybridizing to a first segment of the primary strand of the target nucleotide sequence. The second probe (b) is capable of hybridizing to a second segment of the primary strand of the target nucleotide sequence. The 5' end of the first segment of the primary strand of the target is positioned relative to the 3' end of the second segment of the primary strand of the target to enable joining of the 3' end of the first probe to the 5' end of the second probe, when  
40 the probes are hybridized to the primary strand of the target nucleotide sequence. The third probe (A') is capable of hybridizing to the first probe, and the fourth probe (B') is capable of hybridizing to the second probe (B). The hybridized probes are ligated to form reorganized fused probe sequences. Then, the DNA in the sample is denatured to separate ligated probes from sample DNA. Successive cycles wherein the ligated probes and target DNA undergo the above-described process are performed to increase the amount of detectable DNA in the sample. The amount of cycles performed is dependent upon the sequence used and the sensitivity required of the test. Usually, the cycle can be repeated from 15 to 60 times. At least one of the probes can be conjugated to a signal generating compound.

If the four probes are conjugated to appropriate binding members, the detection of amplified product can be accomplished using standard manual or automated immunoassay procedures known to those skilled in the art. These  
50 procedures include, for example, immunochromatography, ELISA, EIA and MEIA. Hybridization also can be accomplished by following standard dot-, slot- or replica-blot procedures which are known to those in the art. The sequences can be labelled with an appropriate signal generating compound (label), which is capable of generating a measureable signal detectable by external means. The various signal generating compounds contemplated include chromogens, catalysts such as enzymes, luminescent compounds such as fluoroscein and rhodamine, chemiluminescent compounds, radioactive elements such as  $^{32}\text{P}$ , and other labels known to those of ordinary skill in the art. The selection of  
55 a particular label is not critical, but it will be capable of producing a signal either by itself or in conjunction with one or more additional substances. A variety of different indicator reagents can be formed of label and specific binding member. Either the label or a specific binding member can be varied. Examples of specific binding members which

# EP 0 477 972 B1

can be used as a member of the indicator reagent include antibodies, both monoclonal, polyclonal, and fragments thereof; avidin or biotin, biotin and anti-biotin, a carbohydrate or a lectin, a complementary nucleotide sequence, an effector or a receptor molecule, an enzyme cofactor or an enzyme; an enzyme inhibitor or an enzyme; also any antigenic substances, haptens, antibodies, and combinations thereof.

The test sample can be any biological material suspected of containing HPV. Thus, the test sample can be human body tissue, or a test sample which contains cells suspected of containing HPV.

The invention will now be described by way of Examples, which are meant to describe, but not to limit, the spirit and scope of the invention.

The following terms used in the examples are trademarks, tradenames or chemical abbreviations as specified.

TRIS - chemical abbreviation for [tris(hydroxymethyl)aminomethane], used as a buffer.

EDTA - chemical abbreviation for ethylenediaminetetraacetic acid, a chelating agent.

FITC - chemical abbreviation for fluorescein isothiocyanate, a fluorescent hapten derivative.

NHS-ester - chemical abbreviation for N-hydroxysuccinamide ester

MES - chemical abbreviation for [2-(N-morpholino)ethanesulfonic acid], a buffer.

TWEEN®-20 - trademark of Atlas Chemical for polyoxyethylene sorbitan monolaurate, a detergent.

BIS-TRIS - chemical abbreviation for [bis-(2-hydroxyethyl)-amino]tris-(hydroxymethyl)methane, a buffer.

TRITON X- 100® - trademark of Rohm & Haas for nonaethylene glycol octylphenol ether, a detergent.

IMx® - trademark of Abbott Laboratories for an automated instrument for performing microparticle enzyme immunoassay (MEIA).

## EXAMPLES

### EXAMPLE 1

PCR was performed essentially following the package insert of the commercially available Gene-Amp™ kit (document No. 55635-6/89, available from Perkin-Elmer/Cetus, Emeryville, CA). The following reagents were mixed in a 0.5 mL polypropylene tube and used in performing PCR:

Reagent	Final Concentration
Water	(to give final volume = 50 or 100 µL)
Reaction Buffer	10 mM TRIS pH 8.3
	50 mM KCl
	1.5 mM MgCl <sub>2</sub>
	0.01% gelatin
dNTP mixture	200 µM each of dATP, dCTP, dGTP, and TTP
pCR1	1 µM
pCR2	1 µM
plasmid	10 µL 1 ng/100µL
(or control-human placental DNA (Pooled Placental DNA, catalog D-3287, Sigma Chemical Co, St. Louis MO).	
DNA polymerase, <i>Thermus Aquaticus</i>	25 or 63.9 units/1 mL

After mixing, the reaction mixture was overlaid with 100 µL of mineral oil. The tube then was placed in an instrument capable of incubation at several temperatures, and subjected to 30 or 40 cycles of programmed temperature change. The precise cycle of temperature change used, and the instrument used, varied with the experiment, and is detailed in the descriptions of the figures in Example 3.

### EXAMPLE 2

Following the procedure of Example 1, the following sequences were found to amplify sections of papilloma virus types 6, 11, 16, 18, 31, 33, and 61 using PCR.

PCR1: CAGATGTCTC TGTGGCGGCC TAGTG (ID No. 1)

# EP 0 477 972 B1

PCR5: AGGTGTCAGG AAAACCAAAT TTATT (ID No. 5)

PCR14: GAATTAGTTA GACCAITTTAA AAG (ID No. 6)

PCR15: GGGGAAACAC CAGAATGGAT A (ID No. 7)

IWDO: ATCATATGCC CACTGTACCA T (ID No. 8)

Sequence IWDO is derived from a sequence disclosed in International application number PCT/US86/00629 (WO 86/05816).

TABLE 1 shows the sequences and where they map to in the various types.

TABLE 1

SEQUENCES WHICH CAN BE USED AS PROBES OR PCR PRIMERS

PROBE	SEQ	SEQUENCE	SENSE	MAPS TO:	MAPS TO:	MAPS TO:	MAPS TO:	MAPS TO:	MAPS TO:
ID No.				(type 6)	(type 11)	(type 16)	(type 18)	(type 31)	(type 33)
PCR1:	1	CAGATGTCCTGTGTGGCGGCTAGTG	+	5786-5810	5768-5792	5634-5658	5610-5634	5550-5574	5591-5615
PCR2:	2	CGTTTTCATATTTTTTGCAGATG	+	5767-5791	5749-5773	615-5639	5591-5615	5531-5555	5572-5596
OPCR3:	3	AAGTTGTAAGCACCGATGAATATGT	+	5844-5868	5826-5850	695-5719	5671-5695	5611-5635	5652-5676
PCR4:	4	AATGTACCTAAATACCTATATTG	-	6008-5984	5990-5966	865-5841	5841-5817	5784-5760	5825-5801
PCR5:	5	AGGTGTCAGGAAAACCAATTTATT	-	6044-6020	6026-6002	5901-5877	5877-5853	5820-5796	5861-5837
PCR14:	6	GAATTAQTAGACCATTTAAAG	+	1495-1517	1495-1517	1524-1546	1595-1617	1462-1484	1518-1540
PCR15:	7	GGGGAAACACCAGAATGGATA	+	1834-1854	1834-1854	1863-1883	1934-1954	1801-1821	1857-1877
SIWDO:	8	ATCATATGCCCACTGTACCAT	-	1931-1911	1931-1911	1960-1940	2031-2011	1898-1878	1954-1934

note: PCR2, PCR3 and PCR4 are not probes or PCR primers of the invention

## EXAMPLE 3

Linearized plasmids containing full-length papilloma virus inserts in pGEM3 were used as targets. These were pHPV6.1 (HPV6), pSP65.11.5 (HPV 11), p65.16.8 (HPV16), pHPV18H(HPV18), pG3 HPV31 (HPV31), pLNK322,HPV33 (HPV33), and pBR322,HPV61 (HPV61). The Programmable Cyclic Reactor™ (available from Eri-comp, San Diego) was used as the incubation instrument. Following PCR procedures as described in Example 1, 10 μL aliquots were analyzed by electrophoresis through agarose (comprising a 5:3 ratio of NuSieve® SeaKem® GTG, available from the FMC Corp., Rockland, ME) in a buffer comprising 0.089 M TRIS, 0.089 M borate, 2 mM EDTA, and 0.5 ppt ethidium bromide.

FIG. 1 is a photograph of an ethidium bromide-stained 1.2% agarose gel showing results using 63.9 units/mL DNA polymerase, in the DNA Thermal Cycler™ (Perkin-Elmer/CETUS, Emeryville, CA). The samples were heated for 5 minutes at 94°C, then subjected to 40 cycles of a temperature program of: 1 minute at 94°C, 2 minutes at 40°C, and 1.5 minutes at 72°C. The PCR primers used in this case were PCR 1 and PCR5 of Example 2. Examination of the gel following electrophoresis showed bands at the expected positions, i.e. 292 bp. Lane 1, HPV6; lane 2, HPV 11; lane 3, HPV16; lane 4, HPV 18; lane 5, HPV31; lane 6, HPV33; lane 7, HPV61; lane 8, pooled human placental DNA (suspected of having HPV infection); lane 9, molecular weight markers-Hae III digest of ΦX174.

FIG. 2 is a photograph of an ethidium bromide-stained 4% agarose gel showing results using 25 units/mL DNA polymerase, in the Programmable Cyclic Reactor™ (Ericomp, San Diego, CA). Samples in this case were subjected to 30 cycles of a temperature program of: 50°C for one (1) minute, 72°C for two (2) minutes and 95°C for one (1) minute. In this case, the primers PCR1, PCR2, PCR3, PCR4 and PCR5 of Example 2 were used to amplify plasmid

# EP 0 477 972 B1

p65,16,8(HPV 16). Examination of the gel of Figure 2 shows bands at the expected positions, i.e., PCR 1 and PCR4, 235 bp, lane 2; PCR1 and PCR5, 267 bp, lane 4; PCR2 and PCR4, 254 bp, lane 6; PCR2 and PCR5, 286 bp, lane 8; PCR3 and PCR4, 174 bp, lane 10; PCR3 and PCR5, 206 bp, lane 12; molecular weight marker, 123, 246, 369, 492,... bp ladder, lane 1. Note footnote to Table 1.

FIG. 3 is a photograph of an ethidium bromide-stained 1.2% agarose gel showing results using the same conditions as FIG. 1. In this case, PCR14 and PCR15 were used as primers in conjunction with IWDO. The expected size of the amplified PCR product of PCR 14 and IWDO is 437 bp for all of the HPV types tested. The expected size of the product of PCR 15 and IWDO is 98 bp. Products of these sizes appear in the gels, confirming that PCR14 and PCR15, used in conjunction with IWDO, will amplify HPV DNA of types 6, 11, 16, 18, 31, 33, and 61. Lane 1, Molecular weight marker (Hae III digest of FX 174); PCR 14 + IWDO, lanes 2-9: lane 2, HPV6; lane 3, HPV 11; lane 4, HPV16; lane 5, HPV18; lane 6, HPV31; lane 7, HPV33; lane 8, HPV61; lane 9, human placental DNA suspected of being infected with HPV; PCR 5 + IWDO, lanes 10-17: lane 10, HPV6; lane 11, HPV 11; lane 12, HPV16; lane 13, HPV18; lane 14, HPV31; lane 15, HPV33; lane 16, HPV61; lane 17, human placental DNA suspected of being infected with HPV; lane 18, molecular weight marker (Hae III digest of FX174 and Hind III digest of 1 DNA).

## EXAMPLE 4

The following reagents were mixed in a 0.5 mL polypropylene tube as follows for the Ligase Chain Reaction (LCR):

Reagent	Volume	Final Concentration
Water	21 $\mu$ L	
Reaction Buffer	10 $\mu$ L	50 mM EPPS pH7.8 10 mM $\text{NH}_4\text{Cl}$ 10 mM $\text{MgCl}_2$  100 mM $\text{K}^+$ (from all sources) 0.001% BSA 1 mM DDT
Nicotine Adenine Dinucleotide (NAD)	0.5 $\mu$ L	100 $\mu$ L
ProbeA (sense)	4 $\mu$ L	$5.0 \times 10^{11}$ molecules
ProbeA' (antisense, 5'-phosphate)	4 $\mu$ L	$7.5 \times 10^{11}$ molecules
ProbeB (sense, 5'-phosphate)	4 $\mu$ L	$7.5 \times 10^{11}$ molecules
Probe B' (antisense)	4 $\mu$ L	$5.0 \times 10^{11}$ molecules
Target (including human placental carrier DNA at 10 $\mu$ g/mL)	1.5 $\mu$ L	15 ng/50 $\mu$ L
DNA ligase, <i>Thermus therophilus</i>	1 $\mu$ L	

This reaction mixture was overlayed with 30  $\mu$ L of mineral oil. The tube was placed in an instrument capable of incubation at several temperatures (e.g. thermal cycler from Coy Laboratory Products (Ann Arbor, MI) or the Programmable Cycler Reactor™ (available from Ericomp, San Diego, CA), and then subjected to several cycles of programmed temperature change. Each cycle involved incubation at 50°C for one minute and 85°C for one minute.

## EXAMPLE 5

The following procedure was used when performing the Ligase Chain Reaction (LCR), which is described in published European Patent Application No. 0 320 308 A2. The reagents of Example 4 were utilized in the procedure as follows: Two probes (A and B) complementary to immediately adjacent to regions of a target sequence were hybridized and ligated. This ligated probe was denatured away from the target, and hybridized with two additional probes (A' and B') of sense opposite to the initial probes (A and B). The secondary probes then were ligated. Subsequent cycles of denaturation/hybridization/ligation created the formation of double-length probes of both + and - sense.

## EXAMPLE 8

The following sequences were determined to be specific for a portion of the E6 region of HPV type 16:

	<u>Probe</u>	<u>SEQ ID No.</u>	<u>Sequence</u>	<u>Maps to:</u>
	LCR5A	81	GCTGCAAACA ACTATACATG ATATAA	157 - 182
5	LCR5A'	82	pTTATATCATG TATAGTTGTT TGCAGC	182 - 157
	LCR5B	83	pTATTAGAATG TGTGTACTGC AAGCA	183 - 208
	LCR5B'	84	TGCTTGCACT ACACACATTC TAATA	208 - 157

## EXAMPLE 9

Base-denatured plasmids which contained full-length papilloma virus inserts in pGEM3 were used as targets. These plasmids were pG3HPV6(+) (HPV6), pSP 65. 11.5 (HPV11), pSP65.168 (HPV16), p63HPV18H(-)(HPV18), p63:HPV31 (HPV31), pLNK322:HPV33 (HPV33), pBR322:HPV35 (HPV35), pUC19:HPV52 (HPV52), pLNK322:HPV58 (HPV58), pUC9:HPV59 (HPV59) and pBR322:HPV61 (HPV61). All of the oligonucleotides used as probes from Example 8 had chemical labels covalently attached at the ends distal from ligation. These labels were: 5'-fluorescein-LCR5A, 3'-fluorescein-LCR5A', 3'-biotin-LCR5B and 5'-biotin-LCR5B'. Covalent attachment was performed by known methods, i.e., reaction of amine-terminated oligonucleotides with FITC or biotin-NHS-ester essentially following the procedures of Kansal et al., Tet. Letters 29:5537-5540 (1988). The thermal cycler used was obtained from Coy Laboratory Products, Ann Arbor, MI.

Following the LCR procedure of Examples 4 and 5, the mixtures were analyzed using a prototype version of the IM<sub>x</sub> instrument (Abbott Laboratories, Abbott Park, IL), following the protocol for microparticle enzyme immunoassays as follows: A 40μL aliquot of an LCR mixture was diluted 1:1 with distilled water. This diluted mixture was incubated with 50μL anti-fluorescein-conjugated polystyrene microparticles for five (5) minutes to form a suspension of immune complexes on the microparticles. This suspension then was transferred to an inert glass fiber matrix, to which the microparticles became attached. The matrix was washed with buffer (0.3M NaCl, 10 mM TRIS pH8, 0.1%NaN<sub>3</sub>). Any immune complexes attached to the glass matrix was detected by using alkaline phosphatase-labeled conjugate that catalyzed the hydrolysis of 4-methylumbelliferone. The rate at which the 4-methylumbelliferone was generated on the matrix was proportional to the concentration of LCR product formed in the reaction mixture.

Referring to FIG. 4, the graph shows the results obtained from performing LCR on 10<sup>7</sup> molecules of the targets in shown. The rate shown is the rate of generation of 4-methylumbelliferone, and is expressed as fluorescence counts/second/second. Background signal is approximately 10 c/s/s, as shown by the amplification of human placental DNA. The only values above background are those for sample containing HPV16, and those values are about 60 times background signal.

## EXAMPLE 10

The following sequences were determined to be specific for a portion of the E6 region of HPV type 18:

	<u>Probe</u>	<u>SEQ ID No.</u>	<u>Sequence</u>	<u>Maps to:</u>
	LCR6A	85	CTTCACTGCA AGACATACAA ATAA	172 - 195
	LCR6A'	86	pTTATTTCTAT GTCTTGCACT GAA	195 - 173
45	LCR6B	87	pCCTGTGTATA TTGCAAGACA GTAT	196 - 219
	LCR6B'	88	TACTGTCTTG CAATATACAC AGG	218 - 196

## EXAMPLE 11

Plasmids which contained full-length papilloma virus inserts in pGEM3 were used as targets. The plasmids used were those described in Example 9. All of the oligonucleotides used as probes obtained from Example 10 had chemical labels covalently attached at the ends distal from ligation. The thermal cycler was obtained from Coy Laboratory Products, Ann Arbor, MI.

Following LCR procedure described in Examples 4 and 5, the mixtures were analyzed as described in Example 9 using the prototype version of the IM<sub>x</sub> instrument (Abbott Laboratories, Abbott Park, IL).

Referring to FIG. 5, the graph displays the results obtained from performing LCR on 10<sup>7</sup> molecules of the targets. The rate shown is the rate of generation of 4-methylumbelliferone, and is expressed as fluorescence counts/second/



# EP 0 477 972 B1

second. Background signal is approximately 15 c/s/s, as shown by the amplification of human placental DNA. The only values above background are those for sample containing HPV 8, and those values are about 40 times background signal.

## 5 EXAMPLE 12

The following sequences were determined to be specific for a portion of the E6 region of HPV type 18:

10	<u>Probe</u>	<u>SEQ ID No.</u>	<u>Sequence</u>	<u>Maps to:</u>
	LCR7A	89	TATATTGCAA GACAGTATTG GAAC	200 - 223
	LCR7A'	90	pGTTCCAATAC TGTCTTGCAA TTTA	223 - 200
	LCR7B	91	pTTACAGAGGT ATTTGAATTT GCATT	224 - 249
15	LCR7B'	92	AATGCAAATT CAAATACCTC TGTA	249 - 224

## EXAMPLE 13

20 Plasmids which contained full-length papilloma virus inserts in pGEM3 were used as targets. The plasmids were those of Example 9. All of the oligonucleotides from Example 12 which were used as probes had chemical labels covalently attached at the ends distal from ligation. The thermal cyclor was as described in Example 11.

Following the LCR procedure of Examples 4 and 5, the mixtures were analyzed as described in Example 9 using the prototype version of the IMx instrument (Abbott Laboratories, Abbott Park, IL).

25 Referring to FIG. 6, the graph shows the results obtained from performing LCR on  $10^7$  molecules of the targets. The rate shown is the rate of generation of 4-methylumbelliferone, and is expressed as fluorescence counts/second/second. Background signal is approximately 15 c/s/s, as shown by the amplification of human placental DNA. The only values above background are those for sample containing HPV 18, and those values are about 80 times background signal.

## 30 EXAMPLE 14

The following sequences were determined to be specific for a portion of the E6 region of HPV type 16.

35	<u>Probe</u>	<u>SEQ ID No.</u>	<u>Sequence</u>	<u>Maps to:</u>
	LCR8A	93	GTATGGAACA ACATTAGAAC AGCA	352 - 375
	LCR8A'	94	pTGCTGTTCTA ATGTTGTTCC ATAC	375 - 352
40	LCR8B	95	pATACAACAAA CCGTTGTGTG ATTT	376 - 399
	LCR8B'	96	AAATCACACA ACGGTTTGTG GTAT	399 - 376

## 45 EXAMPLE 15

Plasmids which contained full-length papilloma virus inserts in pGEM3 were used as targets. All of the oligonucleotides from Example 14 used as probes had chemical labels covalently attached at the ends distal from ligation. The thermal cyclor was as described in Example 11.

50 Following LCR procedure of Examples 4 and 5, the mixtures were analyzed as described in Example 9 using the prototype version of the IMx instrument (Abbott Laboratories, Abbott Park, IL).

Referring to FIG. 7, the graph details the results obtained from performing LCR on  $10^7$  molecules of the targets. The rate shown is the rate of generation of 4-methylumbelliferone, and is expressed as fluorescence counts/second/second. Background signal is approximately 10 c/s/s, as shown by the amplification of human placental DNA. The only values above background are those for sample containing HPV 16, and those values are about 36 times background signal.

EXAMPLE 16

The attached Appendix (example 16) discloses the sequences of the invention aligned to known sequences.

5 EXAMPLE 16

APPENDIX

HUMAN PAPILLOMA VIRUS

10

ALIGNMENT of TYPES 6, 11, 16, 18, 31, and 33; with CONSENSUS SEQUENCE

15 The appendix lists the sequences of HPV types 6, 11, 16, 18, 31, and 33. It also shows where the sequences of this invention line up with respect to these HPV sequences. In addition, the appendix shows where other sequences, known to the Inventors as of 28 September 1990, and claimed or disclosed by or unknown to others, line up with respect to these sequences.

1. *Sequences and Regions Claimed by Us;*

20 PCR = Sequences per examples 1 through 3 (only PCR1, PCR5 PCR14 and PCR15)

LCR = Sequences per examples 4 through 14 only

2. *Sequences and Regions Unknown to Others and Not Claimed by Us;*

25

PCR = Sequences designated PCR other than those above JJ

LCR = Sequences designated LCR other than those above

3. *Sequences and Regions Claimed by Others;*

(*Italics represents antisense sequences*)

AUS = International application number (Australians) PCT/AU88/00047 (WO 88/06634)

35 WL = International application number (Wayne Lancaster, Wayne State University) PCT/US86/00629  
(WO 86/05816)

BE = European Patent Application (Belgians) 89.033834 (X= T or U)

40 C = International application number (CETUS) PCT/US89/03747 (WO 90/C2821)

O = International application number (Oncor) PCT/US89/O1318 (WO 89/09940)

and

45

4. *Sequences and Regions Disclosed by Others.*

S = Sarkar, F.H. and Crissman, J.D. *Biotechniques* 9 180-184 (1990) (*Italics represents antisense sequences*)

50

55

6 1 gTTAATAACAATCTTgTTTAA AAAATAGGAGGG ACCGAAA ACGGTTCAACCGAAAA  
11 1 cTTAATAACAATCTTAGTTTAA AAAAgAGGAGGG ACCGAAA ACGGTTCAACCGAAAA  
5 33 1 gtaaACTATAATgCcaAGTTTA AAAA AGtAGGGtGTAAACCGAAA gCGGTTCAACCGAAAA  
16 1 actACAATAAT tcAtGTATA AAA ctaAGGGcGTAACCGAAA tCGGTTGAACCGAAAA  
31 1 TAATA ATAATAAT ctTAGTATA AAA AAgTAGGGAGTgACCGAAA GtGgTGAACCGAAAA  
10 18 .1 atTAATActTttaAcaattgTAGTATATAAA AA AGGGAGTaaACCGAAAAcGgtcgGgACCGAAAA  
con --taatatata-ta-aa-tccttag-T-tA-AAAAaag-AGGGagtaACCGAAA-acggttt-aACCGAAAA  
C4-GCCAASTTGGCTTTT  
C5-GCCAGCCCTGGCTTTT  
15 C36-CGGTTSAAACCGAAAA  
C37-CGGTCGGGACCGAAAA  
C38-CGGTTSAAACCGAAAM  
C39-CGGTTCAACCGAAAM  
O15-ATTAATACTTTTAACAATGTAGTATATAAA AA AGGGAGTAACCGAAAACGGTCGGGACCGAAAA-O15  
O24- ACTACAATAAT TCATGTATA AAA CTAAGGGCGTAACCGAAA TCGGTTGAACCGAAAC-O24  
20 S1-CGGTCGGGACCGAAAA  
S3-ACCGAAAC  
6 58 CGGTgTATATAAA CCAGCCCTAAAAaTTAGCAaACGAGGCATTATGAAAAGTgcAaATGCCTCCAC  
11 58 CGGTtATATATAAA CCAGCCCCAAAAaTTAGCAGACGAGGCATTATGAAAAGTaaAGATGCCTCCAC  
25 33 62 CGGTgcATATATAAAGCA aACATTTTgcagtaAGgtActGCACgACTATGTTTCAAGACaCtgAGGA  
16 58 CGGTtAGTATA AAAGCA gACATTTTATGcaCcaAaAgAGAACTGCaATGTTTCCacAGGA  
31 60 CGGTgGTATATAAAGCAcataGtATtTGtgCaAAaccTACAgacGCCATGTTTcaAAAATCctgCAGA  
30 18 66 CGGT GTATATAAA agatgtGagaaacacaCcAcaaTACtatgGCgcgCTTtgAggATCCaaCacg  
con CGGTt-gtatataAagcag--ca-a--at--gcaaaca-agcatt-cgatgttt-aagAtcC--c-ga  
GCC-C4 AUS1-ATGCCTCCAC  
GCC-C5  
35 CGG-C36 C67-AAATCCTGCAGA  
CGG-C37 C68-CCTACAGACGCCATGTTCA-C68  
CGG-C38 C71-GCAGTAAGGTACTGCAC-C71  
CGG-C39 O10-GGATCCAACACG  
O15-CGGT GTATATAAA AGATGTGAGAAACACACCACAATACTATGGCGCGCTTTGAGGATCC-O15  
O24-CGGTTAGTATA AAAGCA GACATTTTATGCACCAAAAGAGAAGTGAATGTTTCCACAGGA(-O24)  
CGGTG-S1 S2-CCGCGCGAAACTCCTAGGTTGTGC-S2  
40 CGGTtAGTATA AAAGC-S3

6 125 GTCTGCAACgacCATAGACCAGTTGTGCAAGACGTTTAATCTaTCTaTGCatACgtTGCAAAATTaAtT  
 11 125 GTCTGCAACAtCtATAGACCAGTTGTGCAAGACGTTTAATCTtTCTtTGcACACtctGCAAAATTCAGT  
 5 33 129 aaaACCAcGAACaTgCATgAtTTGTGcCAAGCaTTGgAgACAACtATACACAACAtgAAcTACAGT  
 16 124 gcGACCCaGAAAgTTaCcAcAgTTATGCaCAGAgcTGcAAACAACtATACATGAtATAAtATTAgAaT  
 31 128 aaGACCTcGgAAaTTgCATGAaCTAaGCTCGGcAtTGgAAAtAcCctacgATGAacTAAGATTgAaT  
 10 18 131 gcGACCctacAAgcTaCctGAtCTgtGCaCGGaAcTGaAcActtCactgcAaGAcATAgAaTaAcctT  
 con g-gacCaagaa--tTacat-AgtTgtGCa-ggc-tTgaa-a-atCtatgcAt-a-aTa-aAaTaaa-T  
 GTCTGCAAC-AUS1 AUS7-GCAAGACGTTTAATCT-AUS7  
 AAGACCTC-C67 C74-ACACTCTGCAAAATTCAGT  
 010-GCGACCCCTACAAGCTACCTGATCTGTGCAAGCACTGAACACTTCACTGCAAGACATAGAAAATACCT-010  
 15 024-GCGACCCAGAAAGTTACCACAGTTATGCACAGAGCTGCAACAACtATACATGATATAATATTAGAAT-024  
 S4-CTGGGTCTTTCAATGGTGTCAATA-S4  
 6 193 GtGTGTTTTGCAaGAATGCAGTACCACaGCAGAGATtTATtCATATGCaTATAAacACCTAAAGGTc  
 11 193 GCGTGTtTTGCAgGAATGCAGTACCACcGCAGAGATATATGCATATGccTATAAgaACCTAAAGGTT  
 20 33 197 GCGTGgAaTGCAaaAAACCTtTGCAaCGAtCTGAGGTATATCAtTTTGCaTTTgCaGATTtTAaCaGTT  
 16 192 GTGTGTACTGCAAgcAACAGTTACTgCGAcgTGAGGTATATGAcTTTGctTTTcggGATTtTgCATA  
 31 196 GTGTcTACTGCAAggtCAGTTAacAgaAACACAGGTATTaGAtTTTGcATTtAcAGATTtAaCaATA  
 25 18 199 GTGTATAtTGCAAgacagtaTTggaActtACAGAGGTATTtGAaTTTGcATTtAaAGATTtAttgTg  
 con GtGTgtatTGCAagaa--catTgacac-a-caGAGGTaTatgaaTtTGCaTtTaaagAttTAa--gT-  
 AUS2-TACGTGACTGGTGGCCGTCTC-AUS2 C73-ACACCTAAAGGTC  
 GC-C74 AUS3-TGAGGTATATGACTTTGCTTTT-AUS3  
 30 C60-GAGGTATWTGAHTTTGc-C60 01-CTAAAGGTT  
 C61-GAGATWTATKCATATGC-C61 02-CTAAAGGTT  
 C69-ACAGTATTGGAACCTACAG-C69 04-GATTTCcAA  
 C70-CAACAGTTACTGCGACG-C70 06-TTATGCATA  
 C72-GACAGTATTGGAACCTACAG-C70 07-TTATGCATA  
 S5-GTGTtTTGCAAGGAATGCAGTACCcA-S5 08-AATACGTAT  
 35 010-GTGTATATTGCAAGACAGTATTGCAACTTACAGAGGTATTtGAATTtGCATTtAAAGATTtATTGTG-010  
 011-TTATTtGTG  
 012-TTATTtGTG  
 013-AATAAACAC  
 017-CTAAAGGTC  
 018-CTAAAGGTC  
 40 020-GATTTCcAG  
 024-GTGTGTACTGCAAGCAACAGTTACTGCGACGTGAGGTATATGACTTTGCTTTTCGGGATTtATGCATA-024  
 025-TTATTtGTG  
 45  
 50  
 55

[illegible]

```

6 329 ATATAGACACTTTgATTATGCTGgATATGCAaCaACAGTtGAAGAAGAAActAAacAAGAcATcTTag
11 329 ATATAGACACTTTAATTATGCTGcATATGCACtACAGTAGAAGAAGAAACcAAtgAAGATATtTTAa
5 33 333 ATATAGACATTATAATTATtCTGTATATGGAaATACATTAGAACAAacAgttAAaAAACCTTTaaTG
16 328 gTATAGACATTATtGTTATAGTtTGTATGGAACAACATTAGAACAgcaAtacAACAAACCGtTgTGTG
31 332 ATTTAGAtggtATaGATATAGTGTGTATGGAACAACATTAGAaaaaTgACaAACAAAGGtaTATGTG
10 18 335 ATTaAGAcattTATtcAgActcTGTGTATGGAgacACATTgGAAAAcTaActAACActGGgtTATaca
con aTatAGAcattTaTaattAt-cTgt-TATGgAacaACATaGAA-Aa-aaactAACaaag-t-Tat-tg
atatagacatt-JJ1
GTATAGACATTAT-AUS8
15 C50-ATAHSACAYATACSTTGTGTGTMATCCTT-C50
C51-ATAHSACAYATACSTTGTGTGTMATC-C51
C52-ATAHSACAYATACSTTGTGTGTMAT-C52
C53-CTGAGACACATACCTCTGTGTGTAACC-C53
C54-CTGAGACACATACCTCTGTGTGTAA-C54
C55-CTGAGACACATACCTCTGTGTGTA-C55
20 O10-ATTAAGACATTATTCAGACTCTGTGTATGGAGACACATTGGAAAACTAACTAACACTGGGTTATACA-O10
O24-GTATAGACATTATTGTTATAGTTTGTATGGAACAACATTAGAACAGCAATACAACAAACCGTTGTGTG-O24
TATATCTGTGAAATTAATACGAC-S6
6 397 AcGTGcTAATTTCGgTGcTACCTGTGTGCACAAaCCGcTGTGTGAAGTAGAAAA ggTAAaACaTATAcT
25 11 397 AAGTGTTAATTTCGtTGTtACCTGTGTGCACAAgCCGTTGTGTGAaATAGAAAAA cTAAAgCACATATt
33 401 AAaTaTTAATTAGGTGTATTATaTGTCAAaGcCtTTGTGTCTcAAGAAAAaAAcGACATgTGGAT
16 396 ATTTGTTAATTAGGTGTATTAAcTGTCAAaGcCacTGTGTCTGAAGAAAAgCAAAGACATcTGGAc
31 400 ATTTGTTAATTAGGTGTATaAcTGTCAAaGcACCGTTGTGTCCAGAAGAAAAACAAAGACATtTGGAT
30 18 403 ATTTaTTAATaAGGTGccTgcgGTGcCAgAaACCGTTGaaTCCAGcAGAAAAActTAGACAccTtaAT
con AttTgtTAATTaGgTGtat--tgTGtCAaAaaCCgtTGtgTccagaAGAAAAaca-agAcacTct--at
AUS4-AATTAATCCACATAAT-AUS4 AUS9-GATTTATTTG
35 AUS5-TGTCATAACCTTGAATGTCT-AUS5
O10-ATTTATTAATAAGGTGCCGTGCCAGAAACCGTTGAATCCAGCAGAAAACTTAGACACCTTAAT-O10
O24-ATTTGTTAATTAGGTGTATTAACTGTCAAaGcCACTGTGTCTGAAGAAAAgCAAAGACATCTGGAC-O24
40
45
50
55

```

6 464 aaccAAGGCgCGgTTCATAAA gCTAAATgtacGTGGAAGGG TCGcTG  
 11 464 gggAAAGGCaCGcTTCATAAAA CTAAATaaCcaGTGGAAGGG TCGTTG  
 33 469 ttAAAcAAAGATTTCATAATAT TtcGGGTcGTtGGGCAGGGCGcTGTgcGgCgTGTtG  
 16 464 AAAAAGcAAAGATTTCATAATATA aGGGGTcGGTGGACcGGtCGaTGTATGtCtTGTtG  
 31 468 AAAAAGaAACGATTCCACAACATAG GaGGAaGGTGGACaCGaCGtTGCATagCaTGTtG  
 18 471 gAAAAacgACGATTtCACAAcATAGctgggcaactataGAgGccaGtgccattcgTGctgcaaccGagc  
 con aaaaAa--acgatTtCatAA--atag-----ctaaaggacg--tgGgcagggcg--tgcatgget--Gttg  
 TGGTGTATAGA-AUS9 AUS6-AAATGTATAGATTTTATTC-AUS6  
 C65-CAACCGAGC  
 010-GAAAAACGACGATTTCACAACATAGCTGGGCACCTATAGAGGCCAGTGCCATTTCGTGTGCAACCGAGC-010  
 024-AAAAAGCAAGATTTCATAATATA AGGGGTcGGTGGACCGGTcGATGTATGTCTTGTtG-024  
 6 512 CcTACACTGC TGGACAACATGCATG GAAGAcATGT  
 11 512 CtTACACTGC TGGACAACATGCATG GAAGAcTtGT  
 33 528 gaggtcccgACGTAGAGAACTGCactgtgAcgTGTAAAAacgcCATGgagagGACACaagcc  
 16 523 cagatcatcAAGAaCACGTAGAGAAAC CCAGcTGTAa tCATGCATGGAGAtACAC C  
 31 527 GagAAGACCTcCGTactGAAAC CCAagTGTAa aCATGCgTGGAGAAACAC C  
 18 539 acgacaGgaAcGACtcCaacgacgcAgagaacaCaAgtataAtattAaGtaTgcAtggACctaaggc  
 con --ga--gagaagaccacgta--aga-Actgca---ccaggtgtAaaacatgcaTgagagAcacaaggc  
 C64-GAACACGTAGAGAAAC CCAG-C64  
 ACGACAGGA-C65  
 C66-GAGGTCCCGACGTAGAGAA-C66  
 010-ACGACAGGAACGACTCCAACGACGCAGAGAAACACAAGTATAATATTAAGTATGCATGGACCTAAGGC-010  
 024-CAGATCATCAAGAACACGTAGAGAAAC CCAG-024  
 6 547 TACCCTAAAGGA TAtGTAtTAGACCTGCAaCCTCCaGACCCTGTAGGGTTACATTGCTATG  
 11 547 TACCCTAAAGGA TATaGTAcTAGACCTGCAGcCTCCTGACCCTGTAGGGTTACATTGCTATG  
 33 590 aACgTTAAAGGA ATATGTtTTAGA TTTatATCCTGAaCCAaCTGAcCTATACTGTATG  
 16 579 TACaTTGCATGa ATATaTGTtAGA TTTGCAACCaGAGaCAaCTGAtCTCTACTGTtATG  
 31 577 TACgTTGCAAGAC TATgTGTtAGA TTTGCAACcTcGAGgCAaCTGACCTCCACTGTtATG  
 18 607 aACaTTGCAAGACattgtatTgcatTTAGAgccccaaaAtgaaattcCggtTGACCTtCtaTGTcAcG  
 con tAC-tT--AgGAc-----at-tgt-tTAGAcctt---catcc-ga-cCa--tGaccTAcacTG-tAtG  
 BE16-ACCAGAGACAACXGAXCXACXGX-BE16  
 BE18-GXXAGAXXXGCAACCAGAGACAACXGAXCXAC-BE18  
 010-AACATTGCAAGACATTGTATTGCATTAGAGCCCCAAATGAAATTCGGTTGACCTTCTATGTcACG-010  
 C89-G  
 C90-G

EP 0 477 972 B1

6 609 AGCAATTAGTAGACAGCTCAGA AGATGA GGTGGACgAaGTGGACggACAAGAttCACAAACCT  
 11 609 AGCAATTAGaAGACAGCTCAGA AGATGA GGTGGACaAGGTGGACaaACAAGAcGCACAACCT  
 5 33 649 AGCAATTAAgTGACAGCTCAGAtGAGGATGaaGgctTGGACcGGCCAGATGGACAA GCACAACCa  
 16 638 AGCAATTAAaTGACAGCTCAGAgGAGGAGGATGaaATAGAtgGTCCAGCTGGACAA GCAGAACCG  
 31 636 AGCAATTAccCGACAGCTCAGAtGAGGAGGATGtcATAGAcgTCCAGCTGGACAA GCAGAACCG  
 10 18 675 AGCAATTAAgCGACtCagagGaaGaaaAcGATGaaATAGA tggagttaatcatcaacatttAcCaG  
 con AGCAATTAA--GACagctcaGAtga-gAtGAtga-aT-GAc-gg-c-gatggacaagacgcacAaCcg  
 AGCAATTAGWAGAC-C89 BE8-GACGAAGXGGACGGACAAGAXXC-BE8  
 AGCAATTAAARYGAC-C90 BE9-GA---GGXGGACGAAGTGGACGGACAAGATTACAAACC-BE9  
 BE13-XGA---GGXGGACAAGGKGACAAAC-BE13  
 BE14-AGAAGAXGA---GGXGGACAAGGKGACAAACAAGACG-BE14  
 15 BE15-CAGAACCG  
 BE17-ACAAGCAGAACCG  
 C62-CGAAGTGGACGGACAAGAT-C62  
 C63-CAAGGTGGACAAACAAGACG-C63  
 010-AGCAATTAAAGCGACTCAGAGGAAGAAAACGATGAAATAGA TGGAGTTAATCATCAACATTTACCAG  
 20 6 671 TTAaaACAACATTtCCAAATAgTGACCTGTTG CTGTGGATGTGAC AGCAACGTtCGA  
 11 671 TTAACACAACATTACCAAAATAcTGACCTGTTG CTGTGGATGTGAC AGCAACGTtCGA  
 33 714 GcCACAGCtgATTActAcATTGTAACCTGTTGT caCActTGTaAC ACCAcAGTTCGt  
 25 16 703 GACAgAGCCcATTACAATATTGTAACCTTTTGTG CAAGTGTgACT CTACgCTTCGg  
 31 701 GACAcAtCCaATTACAATATCGTtACCTTTTGTGT cAGTGTAAGT CTACACTTCGt  
 18 741 ccCgacgagccgaACcAcAaCGTcACacaatGTTGTgtatgtgtGTAAGTgtgaagCcAgAaTTgag  
 30 con g-cacag--cattaCcA-At-gT-ACctgtTGttgt-ctg--g-TGT-Act-----accAcagTtcg-  
 GACAGAGCCCAx-BE15 BE19-AGXGXGACX-----CXACGCXXCCG  
 GACAGAGCCCA-BE17 BE20-XXGCAAGXGXGACX-----CXACGCXXCCG  
 BE24-XXGXAAGXGXGAAGCCAGAAXXGAG  
 BE25-AXGXGXGXGAAGXGXGAAGCCAGAAXXGAG  
 010-CCCGACGAGCCGAACCACAACGTCACACAATGTTGTGTATGTGTTGTAAGTGTGAAGCCAGAATTGAG-010  
 35  
 40  
 45  
 50  
 55



6 728 CTGGTTGTGcAGTGTcACAGaAacAGACATCAGAgAAgTgCAAcAgCTTcTGTtGGGgAACACTAAAcAT  
 11 728 CTGGTTGTGgAGTGCcACAGAcgGAGACATCAGAcAAcTACAAGAcCTTtTGTcGGGCACACTAAATAT  
 5  
 33 771 TTaTGTGTcaAcACTcACAGcaaGtGACcTaCGAACcaTACAgcAACTacTtATGGGCACAgTgAATAT  
 16 760 TTGTGcGTACAAAGCACACAGTAGACATTGtACTtTGTgAAGAcCTGTTAATGGGCACAcTaGGAAT  
 31 758 TTGTGTGTACAgAGCACACAAGTAGAtATTGcAtLATTGCAAGAGCTGTTAATGGGcTcAtTtGGAAT  
 10  
 18 809 cTagtaGTAGaAAGCtCAGcAGAcGAcCTTCGAgcATTcCAGcAGCTGTTtTgAaCaCccTgtcctT  
 con -Tg--tGTacAgcGcaCagaag-aGAcAttcGaacatTgcAa-AgCTgtT-atGggcaCacTaaa-aT  
 XXG-BE19 BE29-AGCAAGXGACCXACGAACCAAXACA-BE29 C42-CCCGTGTGAYYYDTA  
 XXGXGCGXAC-BE20 C43-CTTGTGGGACAGGAA  
 CXAGX-BE25  
 15 BE30-AGXACAGCAAGXGACCXACGAACCAAXACAGCAACX-BE30  
 010-CTAGTAGTAGAAAGCTCAGCAGACGACCTTCGAGCATTCCAGCAGCTGTTTCTGAACACCCTGTCTCTT-010  
 6 796 aGTGTGTCCCATCTGCGC AC CgAaGcCtTAACAAcGATGGCGGACGATTcAGGTACAGAAAAT  
 11 796 TGTGTGTCCCATCTGCGC AC CaAAACCaTAACAAGGATGGCGGACGATTcAGGTACAGAAAAT  
 20  
 33 839 TGTGTGCCCTAcCTGTGC ACaAcAAAtAAACATCAtCtAcaATGGCcGATcCTGaAGGTACAAAtGgg  
 16 828 TGTGTGCCCTAcCTGTGT CAgAAACcATAATCTACcATGGCTGATCCcAGGTACcAATGGGGAA  
 25  
 31 826 cGTGTGCCCTAcCTGTGT aCtAgAcTGTAA CTACAATGGCTGATCCAGCAGGTACAGATGGGGA  
 18 877 tGTGTGTCCgtgTGTgC atCccagCaGTAAGCaACAATGGCTGATCCAGaAGGTACAGAcGGGGA  
 con tGTGTG-CCcAtcTGTgCtaca-aaacaataatcaaCaAtg---G-t---g---gg---ta-ag-ggat  
 C40-CACACRGGGTAGACRCG-C40 C75-ATGGCKGAYCCTGMAGGTAC-C75  
 C41-CACACAGGCACACACG-C41 C76-ATGGCKGAYGATTcAGGTAC-C76  
 ACACAC-C42 C77-ATGGCKGAYCCTTCAGGTAC-C77  
 30 ACACAC-C43 C81-TACCGMCTRGGACKTCCATG-C81  
 C82-TACCGMCTRCTAAGTCCATG-C82  
 C83-TACCGMCTRGGAAGTCCATG-C83  
 010-TGTGTGTCCGTGCTGTGC ATCCCAGCAGTAAGCAACAATGGCTGATC-010  
 35 6 859 GAGGGGTcTGGGTGTACAGGATGGTTTATGGTAGAAGCtATAGTgcAaCACcCaACAGG TAC  
 11 859 GAGGGGTcGGGGGTGTACAGGATGGTTTATGGTAGAAGCcATAGTAGAGCACActACAGG TAC  
 33 906 GctGGGAtGGGGTGTACTGGtTGGTTTgAGGTAGAAGCaGTcATAGAGAgAaAACAGG aGA  
 40 16 895 GaGGGtACGGGATGTAATGGaTGGTTTATGTAGAgGcTGTAgTgAaAaAAAAACAGG GGA  
 31 891 GGGgACGGGATGcAATGGtTGGTTTATGTAGAAGCaGTAAttGACAgAcAgACAGG GGA  
 18 943 GGGcACGGGtGTAAcGGcTGGTTTATGTACaAGCtaTtgTaGACAAaAaACAGGagatgtaat  
 45 con gagGGgagcGGgTGtA-tGGaTGGTTTta-GTAgAaGcT-TagTagA-aaaaaaACAGG-----a  
 C78-TGTAMWGGMTGGTTTATGT-C78  
 C79-TGTAMWGGMTGGTTTcAGGT-C79  
 C80-TGTAMWGGMTGGTTTATGGT-C80  
 C84-ACATKWCCKACCAAAATACA-C84  
 C85-ACATKWCCKACCAAACTCCA-C85  
 50 C86-ACATKWCCKACCAAAATACCA-C86

6 921 ACAAATATCAGACGATGAGGAtGAGGAGGTGGAGGACAGTGGGTATGACATGGTGGACTTTATTGATG  
 11 921 ACAAATATCAGAAGATGAGGAaGAGGAGGTGGAGGACAGTGGGTATGACATGGTGGACTTTATTGATG  
 33 968 TaATATTtCAGAAGATGAGGAtGAAAcGcaGATCAGAGTGGcagGATTTacTAGAgTTTATAGATG  
 16 957 TgcTATaTCAGAtGACGAGAaGAAAAcGacAGTGATACaGGtGaaGATTGGTAGAtTTTATAGtaa  
 31 951 caacATTTcAGAGGACGAAAAcGAAgACAGtAGTGATActGGGAGGATATGGTtGAcTTTATTGAcA  
 18 1009 atcagaTgacGAGGACGAAAAcG caACAG AcACaGGGtcGGATATGGTaGAtTTTATTGata  
 con a-aaat-tcaGA-GA-GAg-AtGaa-a-g-ggatgAcA-tGGgtagGAtaTggTaGAcTTTATTtGat-  
 6 989 A CAGcaATATTACA CAcAATTCacTGGAAGCACAGGCATTGTTTTAAcAGGCAGGAGGCG  
 11 989 A CAGgcATATTACA CAAAATTctGTGGAAGCACAGGCATTGTTTTAAcAGGCAGGAGGCG  
 33 1036 ATtcTAtgGAAaATAgTATACAGGCAGAcACAGAGGCAGCcCgGGCATTGTTTTAAcACAGGAAGgG  
 16 1025 ATgaTAATGAtTATttaAcACAGGCAGAAcAGAGACAGCACAtGCgTTGTTTAcTGCACAGGAAGcA  
 31 1019 ATtgTAATGTaTAcacAAcCAGGCAGAAgCAGAGACAGCACAGGCATTGTTTCATGCACAGGAAGCg  
 18 1071 cacaaggaacATtttgtgAaCAGGCAGAgctAGAGACAGCACAGGCATTGTTcCATGCgCAGGAgGtc  
 con attataatgcatatataataCAGgcagA--cagaG-cAGCaCagGCAtTGTttaat-c-CAGGA-Gcg  
 6 1048 GAcacCcATTATGCGACTGTGCAGGACCTAAAACGAAAGTATTTAGGtAGTCCATATGTTAGTCCTAT  
 11 1048 GAtGCTCATTATGCGACTGTGCAGGACCTAAAACGAAAGTATTTAGGcAGTCCATATGTaAGTCCTAT  
 33 1104 GAgGATgATTtAaATGCTGTgtGcaCTAAAACGAAAGT TTGCCgc  
 16 1093 aAacAACATagAGATGCaGTaCAGGTTCTAAAACGAAAGT AT TtGGTAGTCCa  
 31 1087 gAggAACATGCAGAgGctGTGCAGGTTCTAAAACGAAAGT ATgTaGGTAGTCCt  
 18 1139 cAcaAtgATGCACaAaGtgtTGCAtGTTtTAAAACGAAAGT ttgcaggaggcagcacaga  
 con gA-gatcATt-agaggctgTgcagGttctAAAACGAAAGTattttagg-agtcca--tgtga-tgcc-t  
 BE1-XAAAACGAAAGX-BE1  
 BE2-AGGACCXAAAACGAAAGXAXXXAG-BE2  
 BE3-AGGXXCXAAAACGAAAGXAXXXGG-BE3  
 BE4-AXGXXXXAAAACGAAAGXAXXCAG-BE4

```

5      6 1116 AAACACtATAGCcgAgGCAGTgGAAAGTGaaATAAGTCCACGaTtGACGCCATTAAACTTACAaGAC
      |||
     11 1116 AAGCAATgTAGCTaATGCAGTaGAAAGTGAGATAAGTCCACGgTtaGACGCCATTAAACTTACAaC
      |||
     33 1151 ATGttcacaaagTgTgCGgGagGacgtTGTtGatCGTgTgcaAacCCgtGtAgAacgtCTAttAaTA
      |||
     16 1146 ctTAGTGATATTAG      TGGaTGtGTaGacaATAATATTAGTCCtaGaTTAAAGCTATATGTa
      |||
     31 1141 tTAAGTGATATTAG      TaGtTGtGTGGATtATAATATTAGTCCACGGTTAAAGCTATATGCa
      |||
10    18 1198 aaAcagtcctATTAGggggagcggtggagGTGGATacagAgtTaAGTCCACGGTTACAAAGaaATATctt
      |||
con    a-aca-tatAttagaggcagtgga-gtGtggatagt--t-taaqtcc--g--taaaagctAta-gta

```

```

15      6 1184 AGCCAAAAAGGTAAAGCGACGGCTGTTTcAAAcCaGGGAActAACGGACAGTGATATGGCTATTCT
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    11 1184 AGCCAAAAAGGTAAAGCGACGGCTGTTgAAACacGGGAAtTAACGGACAGTGATATGGCTATTCT
          { | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    20     33 1219 AAaAtAaAgAatGcAcAtacaGAAAACGaAAATAgaTGagcTAGAAGCAGACGGGATATGGCAATACT
          ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    16 1207 TAGAAAAACAAGTagAgCtGCAAAAAGgAGAtTATTTCAAagcGAAGACAGCGGGTATGGCAATACT
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    31 1202 TAGAAAATaAcAGTAAACAcaGAAAACGaAGACTeTTTTGAActtcCAGACAGCGGGTATGGCAATACT
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
    18 1266 TaaAtAgTgggcagAAAAagGCAAAAAGgcSgCTgTTTacAaTaTCAGAtAcTGGcTATGGCtgTtCT
          ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
25 con -a-aaAaaaq-q-Aaa--aq-aaaa-q-a-aatatttgaacta-caGAcAG-GGaTATGGC-at-CT

```

JJ3-tatggctattct  
C87-ATACCGTTANGA  
C88-ATACCGAYAWGA

```

      6 1252 GAAGTGGAAAGCTGgaacgggAACG CAGGTACAGAAAACA TGGCG
          ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    11 1252 GAAGTGGAAAGCTG cAACG CAGGTAGAGAAAACA TGGCG
          ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
35 33 1287 GAAGTGGAAACT CAGCAGAT GGTA cAACA GGTAG
          ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    16 1275 GAAGTGGAAACT CAGCAGAT G tTACA GGTAG
          ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
    31 1270 GAAGTGGAAAC gCAGCAGAT G gTACA GGTAG
          ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
40 18 1334 GAAGTGGAAgc aaCaCAGATtcaggtaacTACaaatggcgcaacatggcggcfaatgtatGTAG
          ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
con GAAGTGGAA-Ctg----gca-caGataggtagag--ACA-----GtaG
   gaagtggaaagctgnnnnnnccnacagat-JJ]
   CTTCACCT-C87
   CTTCACCT-C88

```

```

6 1295          taCCGGAAAATGG          GGGAGATGGTCAGGAAAaGGA
11 1289 A          CCCGGAAAATGG          GGGAGATGGTCAGGAAAaGGA
5 33 1321 A          AagtcAAAATGG          cgACAC AaActtaaAtGActtaGA
16 1306 A          A          gggcGccatgagactgAAACACcAtgtagtcAgtAtagtGg
31 1301 A          GGAG          CAAC          AAACA          AC
10 18 1396 tggcggcagttacGGAGGctatagaCAACgggggacagagggcAACA          AC
con      a-----ggag-----aacgcaaaatgg-----aga--gaaacacgagatggtcaggaaaggga

6 1329 CACAGGaAGGGACATAGAGGG          GGaGGAACATAcAGAGGCGGAAGCgcccacaaACAGtgtaC
15 11 1323 CACAGGgAGGGACATAGAGGGTgagGgGGTGGAAcATaGAGAGGCGGAAGCagtagacGACAGcaccC
33 1358 atCtAGTGGGGtgGGgGAtGaTtcaGAaGTaAGctGTgagacaaatGtAGaTagctGTGAAA
16 1349 tggAAGTGGGGgtGGttgcagTcagtAcaGTAGTGGaaagtggGGGAGagGgTgtTAGTGAaAGAcAcA
20 31 1317 AttAAGT          tgtaTgGTAGTG          ACGGGA          cAcATAGTGAaAGagAgA
18 1445 A          gcagtgtagacggTacaAGTG          AC          aAtAgcaatAtaGAaAat
con      a-caagtagggacagaga-ggt-agga-g--agtgataga-cgg--gaagca---agtGAaaga-a--

25 6 1391 GgGAGCATGCAGgCACAgCAGGAATAT TgGAATTgtTAAATGTAAaGATtTAC GggCagCATT
11 1391 GaGAGCATGCAGACACAtCAGGAATAT TAGAATTacTAAATGTAAgGATATAC GatCtACATT
30 33 1420          atgttACgttgCAGGAA AT TAGtAATGTCTAcAtAgTAGTAATACAAAGCAAatAT
16 1417 cTatAtgcCaAACACcacttacAA ATATTTTAAATGTaCTAAAACTAGTAATGCAAAgGCAGCaAT
31 1361 aTgAAaCtCCAACAC          GtA ATATATTgcAaGTGTTAAAACTAGCAATGgtAAAGctGCTAT
35 18 1487 gTAAAtCcaCAAtgtaccataGcAcAatTAAaagActTGTTAAAGtaAaCAATAaAACAGGaGCTAT
con      gtgaa--t-caa-c-ca-caggaAtAtattagaaatggt-tAaaaa--ag-aaTaaaaagcagc-aT

40 6 1455 AcTGGTAAGTTTAAAGaATGCTTTGGGCTGTCTTTTaTaGATTTAATTAGGCCATTTAAAGTGATA
11 1455 ACaTGGTAAGTTTAAAGAcTGCTTTGGGCTGTCTTTTgTtGATTTAATTAGGCCATTTAAAGTGATA
33 1478 ATTAtaTAAATTTAAAGAGgcCTATGGaTaAGTTTTTaTgGAATTAGTAAGACCATTTAAAGTGATA
16 1484 GTTAGCaAAATTTAAAGAGTTATACGGgGTgAGTTTTTcaGAATTAGTAAGACCATTTAAAGTAATA
45 31 1422 GTTAGgtAAATTTAAAGaATATATGGtGTAAGTTTTatgGAaCTAaTTAGGCCATTTcAAAGcAATA
18 1555 GTTAGcagtATTTAAAGAcacATATGGgtATcaTTTAcAGAttTAgTTAGaaatTTTAAAGtgATA
con      -ttaggtaaaTTAAAGA-t--TatGGgcT---tTTTataGA-tTA-TtAG-ccaTTTAAAGtgATA
JJ4-ttaggttagaccattttaaagtata

```

6 1523 aAACaACATGTTtaGATTGGGTGGTtaGCAGGgTTTGGTATACATCATAGCATAtCAGAgGCATTTCaA  
 11 1523 gAACcACATGTgCcGATTGGGTGGTtGCAGGATTGGTATACATCATAGCATAGCAGAtGCATTTCaA  
 33 1546 AAACAAGcTGTaCcGATTGGTGTATaaCAGGATaTGGAAATTAgTCCatcagTAGCAGaAAGTTAAAA  
 16 1552 AAtCAACgTGTtgcGATTGGTGTATtGCTGCATTGGAcTTACACCCaGtaTAGCtGACAGTaTAAAA  
 31 1490 AAAgCACaTGTACTgATTGGTGTgTAGCTGCgTTTGGAGTTACAggtACAgTtGCAGAAGGATTAAAA  
 18 1623 AAAcCACgTGTACaGATTGGgtTacAGCTataTTTGGAGTaAacccaACAAaTaGCAGAAGGATTAAAA  
 con aAac-AcaTGTacaGATTGG--t-tagC-ggaTtTGGaaT-aatccta-aaTagCaGAaggatTtaAA  
 aaacaacNtgtNcagattgg-JJ4  
 6 1591 AAaTTAATTGAGCCATTAAGTTTATATGCACATATACAATGGCTaACAAATGCATGGGGAATGGTAtT  
 11 1591 AAgTTAATTGAGCCATTAAGTTTATATGCACATATACAATGGCTtACAAATGCATGGGGAATGGTAcT  
 33 1614 gtATTaATTaAACAgcATAGTTTgTATaCtCATtTACAATGTTTAACTtGcGataGaGGAATaaTaaT  
 16 1620 ACACTATTaCAACaATATTGTTTATtTaCacaTtCAAAGTTTAGCATGTTCaTGGGGAATGGTTgT  
 31 1558 ACcCTATTgCAACCATATTGTTTgTATtGCCATtTaCAAAGTTTAGCATGTTCCtGGGGcATGGTTAT  
 18 1691 ACaCTAAaTaCAGCCATtTataTTaTATgCCATaTtCAAtGTcTAGAcTGTaaaTGGGgagTatTaaT  
 con aca-TAAtTcA-Ccat-tagtTTaTATgcaCAtaTaCAAtGt-Ta-catgtgcatGgGGaaTggTaaT  
 6 1659 gTTAGTATTAtTAAGaTTTAAAGTAAATAaaAGtAGaAGTACCGTtGCACGTACAcTtGcaACGcTAT  
 11 1659 ATTAGTATTAAATAAGgTTTAAAGTAAATAAGAGCAGATGTACCGTgGCACGTACAtTaGGTACGTTAT  
 33 1682 ATTAAtTgTTAAAtAGATTAggTGTAGcAAAAACAGgTtAACAGTaGCaaaaTaATGaGTAAtTTAT  
 16 1688 GTTAcTaTTAGTaAGATaTAAATGTGGAaaaaATAGaAACAATTGAAAAATTgGTGcTAAAcTAT  
 31 1626 GTTAAAtgCTtGTGAGATtTAAATGTGcAAAAAATAGaaTAACAATTGAAAAATTaTTAgaaAAATTAT  
 18 1759 aTTAgccCTgtTGcGtTacAAATGTGgtAAgAgTAGAcTAACAgTTGctAAAaggTtTAagtAcgTTgT  
 con -TTAgtatTa-TaaGaTttAaatgt-gtAAaA-tAGa-taACagTtGcaaaa-tatTaggtA-gtTaT  
 6 1727 TAAATATACCTGAAAAcCaATGTTAATaGAGCCaCCaAAAAATACAAAGTGGtGTtgcAGCCCTGTAT  
 11 1727 TAAATATACCTGAAAAcCaATGTTAATtGAGCCtCctAAAAATACAAAGTGGcGTAcGAGCCCTGTAT  
 33 1750 TATcaATACCTGAAAcaGTATGgTtATAGAGCCaCCAAAAATTACGgAGccaaAaCatGtGCATTGTAT  
 16 1756 TATGTgTgTCTcCAAtgTGTATGaTgATAGAGCCtCCAAAAATTgCGTAGtACAGCAGCAGCATTATAT  
 31 1694 TgTGTaTATCTaCAAAaTTGTATGTTAATTcAGCCACCcAAATTaCGTAGcACAGCTGCAGCATTATAT  
 18 1827 TacacgTAcCTGaAAcTTGTATGTTAATTCAaCCACCaaAAATTgCGaAGtAgTgTGCAGCAcTATAT  
 con TatataTacCTGaAAattgtATGtTaAT-gAgCCaCCaAAAtTaCgaAGtagag--gcaGCa-T-TAT

6 1795 TGGTTTCtGtACAGGtAtaTCAAAATGCCAGTACAGTTATAGGGGaaGCaCCaGAATGGGATAACaCGCCA  
11 1795 TGGTTTAGgACAGGcAtTtCAAAATGCaAGTACAGTTATAGGGGAgGCgCCgGAATGGATAACgCGCCA  
33 1818 TGGTTTAGAACAGcAaTgTCAAAcATTAGTGATGTacAAGGtacaACaCCtGAATGGATAgatAGACT  
16 1824 TGGTATAaaAACAGGtAtaTCAAAATATTAGTGAAgTgTATGGaGAcACgCCAGAATGGATAcAAAGACA  
31 1762 TGGTAcAGAAACAGGAATgTCAAAcATTAGcGATGTATATGGtGaaACACCAGAATGGATAgAAAAGACA  
18 1895 TGGTAtAGAACAGGAATaTCAAAATATTAGtGaaGTAatgGGaGAcACACcTgAGtGGATAcAAAGACT  
con TGGT-tagaACAGgaAtaTCAAAAtattAGtgaaGTaa-aGG-gaaaCaCCaGAaTGGATA-aaaGaCa  
BE32-AXAXCAAAAXXAGXGAAGX-BE32 JJ6-tggataNaaagaca  
6 1863 aACaGTTATTGAACAcgGgTTGGCaGACAGTCagTTTAAATTAACaGAAATGGTGCAGTGGGCgTATG  
11 1863 gACcGTTATTGAACATAGtTTGGCTGACAGTCAATTTAAATTAACtGAAATGGTGCAGTGGGCATATG  
33 1886 AActgTTTTTACAACATAGcTTTAATGATAaTatATTGATtTTAagTGAAATGGTACAGTGGGCATATG  
16 1892 AACAGTATTACAACATAGTTTTTAATGATtgTACATTTGaaTTATCaCagATGGTACAAATGGGCcTAcG  
31 1830 AACAGTATTACAgCATAGTTTTTAATGAcAcaACATTTGATTTGTcCCAAATGGTACAAATGGGCATATG  
18 1963 tACtaTtaTACaAcATgGaaTagATcAtAgcAatTTTGATTTGTcAgAAATGGTACAAATGGGCATtTG  
con aACagTt-TacAaCAtaGttTt-atGA-agtaaaTTTgA-TTa-cagAaATGGTaCA-TGGGcATatG  
aAcNgttatacaacatagtttNgatgat-JJ6  
6 1931 ATAATGAcAtaTgCGAgGAGAGTGAAaATtGCATTTGAATATGCACAAaGgGGAGatTTTGATTCtAAT  
11 1931 ATAATGAtAtTtGtGaaGaaAGTGAGatAGcATTTGAATATGCACAgCGTGGAGAcTTTGAcTCcAAT  
33 1954 ATAACGAgTtAacgGACGATAGTGAcATTGCATATtAtTATGCACAACTTGcAGAtTCaAATAGTAAT  
16 1960 ATAATGAcAtAgTaGACGATAGTGAAATTGCATATAAAATATGCACAATTgGCAGACACTAATAGTAAT  
31 1898 AcAATGAtgTtATgGATGATAGTGAAATTGcCTATAAAATATGCACAATTAGcTgACAGTgATAGTAAT  
18 2031 AtAATGAgcTgAcaGATGAAAGcGATgGCaTtTgAATATGCcctATTAGCaGACAGcaAcAGcAAT  
con AtAATGA-aTaa--GA-GATAGtGAaATtGCAT-TgAaATGcacaatt-GcAGAc--t-AtagtAAT  
6 1999 GCACGaGCaTTTtTAAATAGcAAATATGCAGGCAAAATATGTgAAAGATTGTGCAAcTATGTGtAGACA  
11 1999 GCaaGgGCcTTTtTAAATAGTAAATATGCAGGcCAAAATATGTAAAGATTGTGCAATTATGTGcAGACA  
33 2022 GcTgcTGCaTTTtTAAAAAGTAACtCACaaGCAAAATaGTAAAGGAcTGTGgAATAATGTGTAGACA  
16 2028 GCaaGTGCcTTtTcTAAAAAGTAATTCACAGGCAAAATtGTAAGGATTGTGcAACAAATGTGTAGACA  
31 1966 GCAtGTGCaTTTtTAAAAAGTAATTCgCAGGCAAAATaGtTAAAGATTGTGgAACAAATGTGTAGACA  
18 2099 GCAGcTGCcTTTtTAAAAAGCAATTgCaaGcTAAAtattTaAAAGATTGTGccACAATGTGcAAACA  
con GCa-gtGC-TTTtTAAAAAGTAATtcqCAgGCCAAAA--tgTAAAAgATGTGcaAcaATGTGtAgACA

6 2067 TTATAAACATGCAGAAATGAggAAGATGTCTATAAAACAATGGATaAAacATAGGGGTtCTAAAaTag  
 11 2067 TTATAAACATGCAGAAATGaaaAAGATGTCTATTAAACAATGGATTAAgtATAGGGGTaCTAAAgTtG  
 33 2090 TTATAAAAAGCAGAAAAAcgtAAAAATGTCaATagGACAATGGATAcAAagTAGATGTGAaAAAacAa  
 16 2096 TTATAAACGAGCAGAAAAaaACAAATGagtATGaGtCAATGGATAAAAaTaTAGATGTGAtAggGTAg  
 31 2034 TTATAAACGAGCAGAAAAACGACAAATGtccATGgGACAGTGGATTAAAagTAGATGTGACAAAgtTa  
 18 2167 TTATAggCGAGCccAAAAACGACAAATGaATATGtCACAGTGGATacgAttTAGATGTtcaAAAaTag  
 con TTATAaac-aGcagAAAAa-ga-AaATGtctATyagaCAaTCGATaaaataTAGatGTg-tAaa-tag  
 JJ11-tggataaaaatatagatgtNctsaatatag  
 6 2135 AagGcacAGGaaAtTGGAAaCCAAATTGTaCAaTtCTAcGACATCAAAaTaTAGAATTcATTCCtTTT  
 11 2135 AcaGTGtAGGtAAcTGGAAgCCAAATTGTgCAGTtCTAAAGACATCAAAACATAGAAATTTATTCcATT  
 33 2158 ATGATGGAGGaaATTGGAGaCCAAATaGTaCAGTtGTAAAGATATCAAAACATtGAATTTAcagCATT  
 16 2164 ATGATGGAGGTGATTGGAAaCCAAATtGtTAtGTTTTTAAGgTATCAAggtgTAGAgTTTATGTCATT  
 31 2102 gTGACGaAGGTGAcTCGAGGgAcATAGTAAAGTTTTTAAGATATCAACAAATAGAAaTTTgTGCATT  
 18 2235 aTGaaGggGGaGAtTGGAGaccaATAGTgcAaTtccTgcGATAcCAACAAATAGAgTTTaTaaCATTT  
 con atgatggaGG--AtTGA--ccsAT-GTAcagTtT-TaaGatAtCAaaa-aTaGaaTTtat--CATTT  
 atgatggaggaaattgga-JJ11 JJ12-catTT  
 6 2203 TTAActAAAtTtAAATTATGGCTGCACGGtACGCCaAAAAAAAcTGcATAGCCATaGTAGGcCCtCC  
 11 2203 TTAAGcAAAcTaAAATTATGGCTGCACGGaACGCCcAAAAAAAtTGTATAGCCATTGTAGGgCCACC  
 33 2226 TTAgtGTCATTtAAAAagTTTTTaaAGGTATACCaAAAAAAAgcTGTATgcTAATTTgTGGaCCAGC  
 16 2232 TTAaCTGCATTAAAAgaTTTTTgcAAGGcATACCTAAAAAAAtTGcATaTTAcTATATGGTGCAGC  
 31 2170 TTAaCTGCATTAAAgctgTTTTTAAAGGAgTgCCaAAgAAAAAcTGTATtTTaATaCATGGTGCACc  
 18 2303 TTAaggAGCcTTAAAtcaTTTTTAAAGGAaccCCcAAaAAAAAtTGtTTagTAtTtTgTGGacCAGC  
 con TTAA-tgcatTaAAattaTttT--AaGGaa-gCCaAAaAAAAa-TGtaTagtaaT-t-tGG-cCa-C  
 ttaagtgcattaaaattatttttgcaggNacNccNaaaaaaa-JJ12  
 6 2271 aGAtACTGGGAAaTCGTaCTTTTGTATGAG TTAATaAgcTTTcTaGGaGGtACAGTTATTAGTcAT  
 11 2271 tGAcACTGGGAAGTCGTgCTTTTGcATGAG TTAATtAAGTTTTTgGGgGGaACAGTTATTAGTTAT  
 33 2294 aAAaACAGGaaAGTCATatTTTGGaATGAG TTAATacAGTTTTTaaAAGGGTgTGTtATATcaTgT  
 16 2300 TAAaACAGGTAAATCATtATTTGGtATGAG TTAATGAAaTTTcTgCAAGGGTcTGTAATATgtTtT  
 31 2238 TAATACAGGTAAATCATATTTTGGaATGAGccTTATTGAgCTTTtTACAAGGATgTaTAATATCATaT  
 18 2371 aAATACAGGaaAATCATATTTTGGaATGAGttTTAT acaCTTTaTACAAGGAgcagTAATATCATtT  
 con -aAtACaGG-AAATCaTatTTTgGaATGAG-tTTAataaacTTTtTacaagGatc-gT-ATat--taT

6 2338 GTAAATTCCaGCAGCCATTTtTGGtTgCAaCCgtTAgtAGATGCTAAgGTaGCATTgTTaGATGATGC  
 11 2338 GTtAAATTCCTGCAGCCATTTtTGGtTaCAGCCAcTaaCgGATGCAAAAGTgGCATTaTTgGATGATGC  
 33 2361 GTAAATTCTAAAAGtCacTTTTTGGTTgCAGCCATTAtCAGATGCAAAAATAGGAATGaTAGATGATGt  
 16 2367 GTAAATTCTAAAAGcCATTtTTGGTTACAACCATTAGCAGATGCCAAAATAGGtATGTTAGATGATGC  
 31 2306 GcAAATTCAAAAAGTCATTtTTGGTTACAACCACtTgGctGATGCTAAAATAGGCATGTTAGATGATGC  
 18 2438 GtgAATTCCActAGTCATTtTTGGTTggAACCGtTaaCaGATaCTAAgTgGcCATGTTAGATGATGC  
 con GtaAAATTCaaaAG-CatTTtTGGtT-cAaCCatTagcaGATgCtAAa-TaG-aaTgtTaGATGATGC  
 6 2406 aACACAgCCATGTTGGAtATATATGGATACATATATGAGAAAtTgTTAGATGGTAATCCTATGAGtA  
 11 2406 cACACAACCATGTTGGACATATATGGATACATATATGAGAAAccTaTTAGATGGTAATCCTATGAGcA  
 33 2429 aACgCcAatAaGTTGGACATATATAGATGATtACATGAGAAATGCgTTAGATGGAAATgaAATTTCaA  
 16 2435 TACAgTGCCcTGTGGAAcTAcATAGATGAcAttTAAGAAATGCATTgCATGGAAATtTAGTTCTA  
 31 2374 TACAAGGCCaTGTGGcAtTAtATAGAcAAtTAccTAcGAAATGCACtAGATGGCAACcCTGTATCTA  
 18 2506 aACgACCaCgTGTGGAcATActTtGATAcCTAtaTgaGAAATGCgtTAGATGGCAAtCCaATAagTA  
 con aACaccgccatGTTGGAcATAtaTaGAta--tAtaTgaGAAAtgc-tTaGATGG-AAAtcc-aT---tA  
 JJ15-gttggacatatatNgatacNtatatgagaatgcgttagatgg-JJ15  
 6 2474 TtGAcAGAAAgCATaaAGCATTgACATTAATTAAaTGTCACCCtCTgCTaGTaACgTCCaACaTAGAt  
 11 2474 TAGATAGAAAACATAGAGCATTAACATTAATTAAgTGTCACCCgCTaCTgGTTACaTCAAAATAGAc  
 33 2497 TAGATGTgAAACATAGgGCATTAGTgCAAtTAAATGTCCACCACTgCTtcTTACcTCAAAATACAAAT  
 16 2503 TgGATGTAAAGCATAGAcCATTgGTaCAAcTAAATGcCCTCCATTATTaATTACATCTAACAtTAAT  
 31 2442 TaGATGTAAAGCATAAAgCtTTAATgCAgtTAAATGTCTCCtTTATTgATTACATCTAATATAAAT  
 18 2574 TtGATagAAAGCacAAAcCaTTAATaCAAcTAAATGTCTCCaATAcTAcTaACCaCAATATAcAT  
 con TaGAt--aAAgCATa-agCaTTaatacaa-TaAAaTGtCC-CCacTacTa-TtACatCaAAAtataaAt  
 6 2542 ATTActAAAGAaGAtAAATAtAAgTATTTACATAcTAGAGTaACaACATTtACATTtCCAAATCCATT  
 11 2542 ATTAGCAAAAGAgGAgAAATAcAAATATTTACATAGTAGAGTtACcACATTtACATTtCCAAATCCATT  
 33 2565 GCaGGCACAGAcTCTAGATGGCCATATTACATAGTAGATTaACaGTATTgAATTtAAATCCATT  
 16 2571 GCTGGTACAGATTCTAGgTGGCCtTATTACATAaTAGATTGGTGGTgTTTACATTtCCtAAAgagTT  
 31 2510 GCAGGTAAGGATgAcAGATGGCCATAcTACATAGcAGAcTGGTGGTtTTTACATTtCCAAATcCATT  
 18 2642 cCAGCaAAGGATaAtAGATGGCCATAttTagAaAGtAGaATaacaGTaTTTgaATTtCCAAATgCATT  
 con gc-ggtAaaGAtgatAgaTggccATAttTAcAtAgTAGA-TaacagtaTTTAcATTtCCAAATccaATT



6 2610 CCCctTTTGACAGAAATGGGAATGCAGTgTATGAACTgTCAaATaCAAACTGGAAATGTTTtTTTGAAA  
 11 2610 CCCcTTTGACAGAAATGGGAATGCAGTATATGAACTATCAGATGCAAACTGGAAATGTTTcTTTGAAA  
 5 33 2633 CCCATTtGAtGAAAATGGtAAcCCAGTGTATGCaATAAATGATGAAAAtTGGAAATCCTTTTTCTCAA  
 16 2639 TCCATTtGACGAAAACGGAAATCCAGTGTATGAgcTtAATGATAAgAACTGGAAATCCTTTTTCTCAA  
 31 2578 TCCATTtGACAAAAACGGAAATCCAGTATATGAAtTAAGTGATAAAAACTGGAAATCCTTTTTCTCAA  
 10 18 2710 TCCATTtGAtAAAAAtGGcAATCCAGTATATGAaTAAaTGAcAAAAAtTGGAAATgtTTTTtgaaa  
 con -CCaTTTGACaaaAAAtGG-AAtcCAGT-TATGaacTaaatgAtaaaaAAcTGGAAAT--TtTTT---AA  
 6 2678 GACTGTCTGTCaAGCCTAGACATTcAGGATTcTcGAGGA CGAGGAA GATGGAAGCAATAGCCAA  
 11 2678 GACTGTCTGCCAGCCTAGACATTgAGGATTcAGAGGA CGAGGAA GATGGAAGCAATAGCCAA  
 33 2701 GGACGTGGTgCAaATTAGATTTaataGAGGAAGAGGA CAAGGAAAACcATGGAGGAAATATCagc  
 16 2707 GGACGTGGTcCAGATTAAgTTTGACAGGAGcGAGGA CAAGGAAAACGATGGAGACTCTTTgcCA  
 20 31 2646 GGACGTGGTgCAGATTAAATTTGCACGAGGAAGAGGA CAAAGAAAACGATGGAGACTCTTTcCA  
 18 2778 GGACaTGGTcCAGATTAgATTGACAGGAGAGGAagatgcAGAcAcCGAaGGAaAcCCTTTCgga  
 con GgacgTgTccAgatTAgattTgcacGaggaaGAGGA---c-agGaaaacgAtGGA--ca-T-tcc-a  
 6 2740 GCGTTTAGATGCGTGCCAGGAaCAGTTGTTAGAACTTTATGAAGAAAACAGTAcTGAccTACACAAAC  
 11 2740 GCGTTTAGATGCGTGCCAGGAtCAGTTGTTAGAACTTTATGAAGAAAACAGTAcTGAAtTACACAAAC  
 30 33 2766 ACGTTTAAATGCagtgCAGGAGAAAATACTAGAtCTTTAcGAAGcTGATAaaaACTGAtTACcatcAC  
 16 2772 ACGTTTAAATGTGTGTGCAGGACAAAATACTAacACATTATGAAAATGATAGTACAGAcTACGTGAcC  
 31 2711 ACGTTTAAATGTGTGTGCAGGACAAAATAtTAGAACATTATGAAAATGATAGTAAAcgaCTtGTGAtC  
 35 18 2846 ACGTTTAAgTtgcgTgCAGGACAAAATcaTAGAcCAcTATGAAAATGAcAGTAAAgacaTagacagcC  
 con aCGTTTAAaTgcgTg-CAGGAcAaaaTatTAgaac-tTAtGAA-atgA-AgtAc-gaccTAcacaaaC  
 6 2808 AtgTatTGCATTGGAAATGCATGaGAcAtGAAAGTGTATTAtTAtAtAAAGCAAAACAAATGGGCCTa  
 40 11 2808 AcATTAtGCATTGGAAATGCATACGAtTGGAAAGTGTATTAcTAcAcAAAGCAAAACAAATGGGCCTg  
 33 2834 AaATTGAaCATTGGAAAcTgATACGCaTGGAgTGTCTTTATTgTAtAcAGCCAAACAAATGGGATTT  
 45 16 2840 ATATAGACTATTGGAAACAcAtgCGCCTaGAATGTGCTaTtTatTAcAAgGCCAGAGAAATGGGATTT  
 31 2779 ATATAGACTATTGGAAACAtAttCGAcTtGAATGTGtAtTaaTgTATAAAGCAAGAGAAATGGGAATA  
 18 2914 AaATAcAgTATTGGcAACTaATaCGTtgGAAaATGcAaTatTcTtTgcAGCAAGGGAaAcAtGGcATA  
 50 con AtaTagag-ATTGGaAAc-cATAcGactgGAa-gTG-atTatt-tataaaGCaA-a-AAatgGG--Ta  
 55

6 2876 AGCCACATaGGaATgCAAGTAGTgCCACCATTAAaGTGTCCaAgCaAAAGGACATAATGcCATTGA  
 |||||  
 11 2876 AGCCACATcGGgtTaCAAGTAGTaccACCATTAAcGTGTcAGAGAcTAAAGGACATAATGcCATTGA  
 |||||  
 5 33 2902 tcACATtTatGCCACCagGTGGTGCcttCtTTGttaGcATCAAAGACcAAAGCATTTCaAGtAATTGA  
 |||||  
 16 2908 aaACATATTaACCACCAaGTGGTGCCAaCacTgGcTGTATCAaAGAaTAAAGCATTACaAGCAATTGA  
 |||||  
 31 2847 CAcAgTATTaACCACCAGGTGGTGCCAGcGTTgtCaGTATCAaAGgccAAAGCcTTACaAGCTATTGA  
 |||||  
 10 18 2982 CAgAcattAaACCACCAGGTGGTGCCAGCcTataacaTtTCAAAaagtAAAGCacataAAGCTATTGA  
 |||||  
 con aaccataTaa-ccacCA-GTgGTgCCa-Cattgac-gtaTCaaAgactAAAGCa--t-AaGctATTGA  
 JJ18-tcaaagactaaagcacataaaagcNattga  
 15 6 2944 AATGCAAAATGCATTTAGAATCaTTAttAaggACTgAGTATaGTaTGGAAcCgTGGACATTACaAGaa  
 |||||  
 11 2944 AATGCAAAATGCATTTAGAATCcTTAgcAAAAACTCAGTATgGTgTGGAAcCtTGGACATTACAGGAcA  
 |||||  
 33 2970 ACTaCAAAATGgCaTTAGAgACATTaAgTAAATCACAGTATAGTAcagcCAaTGGACATTgCAACAA  
 |||||  
 20 16 2976 ACTgCAAcTaACGTTAGAAACaATaATAACTCACAATATAGTAATGaaaAgTGGACATTaCAAGAcg  
 |||||  
 31 2915 ACTaCAAAATGATGTtGAAACatTaaATAACACTgAATACAAAAATGAGGAcTGGACaATGCAGcAA  
 |||||  
 18 3050 ACTgCAAAATGgcccTacaAggcccTgcacAaAgTcgATACAAAaccGAGGAtTGGACAcTGCaAgAcA  
 |||||  
 25 con AcTgCAAAATg-c-tTagAaacatTa---aaaactca-TAtagta--gaaca-TGGACATt-CAagA-a  
 actgcaaatgg-JJ18  
 6 3012 CaAGTTATGAAATGTGGCaAACACCACC tAAACGcTGtTTTAAAAAACgGGGcAAaACTGTaGAaGT  
 |||||  
 11 3012 CcAGTTATGAAATGTGGCTAACACCACC CAaACGgTGcTTTAAAAAACAGGGAaAtACTGTgGAGGT  
 |||||  
 30 33 3038 CaAGCTTaGAgGTGTGGCTtTgTaACCACC AaaATGTTTAAAAAACaAGGAgAaACAGTaactGT  
 |||||  
 16 3044 ttAGCCTTGAAGTGTATTtAACTGCACCAAC AGGATGTaTAAAAAACATGGATATACAGTgGAaGT  
 |||||  
 31 2983 CAAGtCTTGAAGTGTATTtAACTGCACCTAC AGGgTGTTTAAAAAACATGGATATACtGTAGAgGT  
 |||||  
 35 18 3118 CAAtGcgagGAACTaTggaatACaGaACCTACTcactgcTTTAAAAAA ggTGGccAaACaGTACaAGT  
 |||||  
 con caaG-t-tGAa-TgTggctaac-gcACCaacaa-g-tgttT-AAAAAacatGGA-A-AC-GTagaAGT  
 40 6 3079 tAAATTTGA TGGCTGTGcAaACAATaCAATGGATtATGTGGTATGGACAgATgTgTatgTGCAGG  
 |||||  
 11 3079 aAAATTTGA TGGCTGTGAAGCAATgtAATGGAgTATGTGGTATGGACAcATATATAccTGCAGG  
 |||||  
 33 3105 GCAATaTGA caatGACAAAaAaAATACAATGGATTATACAAACTGGggtgAaATATATATTataG  
 |||||  
 16 3111 GCAGTTTgATGG aGACATatgCAATACAATGCATTATACAAACTGGAcAcATATATATATTGTG  
 |||||  
 45 31 3050 GCaaTTTgATGG tGAtGTAcACAACcATGCATTATACtAACTGGAAAtTTATATAccTaTGTA  
 |||||  
 18 3185 atAtTTTGATGGcaacaAaGacaAttgtAtgAcctATgtagCatggggacAgtgTgtatTataTgacTg  
 |||||  
 con gcAaTtTGATggcaacgatgaaaacaatacaAtggAttat-caaactggacagatataTataTg--tg  
 50  
 55

6 3144 ACAATGACaCCTGGGTAAAGGTgcaTAGTatgGTAGATGctAAGGGtATATATTACACATGTGGACAA  
 5 11 3144 ACAACGACTCATGGGTAAAAgTaaCTAGTtccGTAGATGCCAAGGGcATATATTATACATGTGGACAA  
 33 3170 AgGAAGAtaCATGTACTATGGTtACaGGGaAGTAGATTATAtaGGTATgTATTATATACATAActGt  
 16 3176 AAGAAGcatCAGtaACTGTGGTAgAGGGtCAAGTTGAcTATtAtGGTtTaTATTATGTtCATGAAGGA  
 10 31 3115 tAGAtGGceaATGtACTGTtGTgGaaGGGcCAAGTTAATTgTAAGGGcAtTtATTATGTACATGAAGGA  
 18 3253 atGcaGgaacATGggacaaaaccGctacctgtGTaAgTcacAgGGGatTgTATTATGTAAAgGAAGGg  
 con a-gaaGacacatgg-cta-ggt-g-t-gt-aaGTagattataagGGTaTaTATTAt-tacatgaagga  
 15 6 3212 TTAAAAACATATTATGTAAAcTTTgtaAAAGAGGCAGAAAAGTATGGgAGcACCAAAcATTGGGAAGT  
 11 3212 TTAAAAACATATTATGTAAATTTTAAaAAAGAGGCACAAAAGTATGGTAGtACCAAtCATTTGGGAAGT  
 33 3238 gaAAAggtATAITTTTaaAtATTTTAAAGAgGATGCTGcAAAGTATtcTAAACAcAAaTgTGGGAAGT  
 20 16 3244 atAcgAACATATTTTGTgCagTTTAAAGAtGATGCAGAAAAATATaGTAAAAaTAAAgTATGGGAAGT  
 31 3183 cAtAtAACATATTTTGTAAAtTTTAcAGAAgAGCAaAAAAATATGGGActggTAAAAaATGGGAAGT  
 18 3321 tAcAacACgTtTTaTaTAgAaTTTAAAgTGAatgtgAAAAATATGGGAacacaggtAcgTGGGAAGT  
 25 con t-taaaaaCaTaTT-TgtaaAtTTTaaa-aaGAggcagaAAA-TATgg-Aa-ac-aaaaa-TGGGAAGT  
 6 3280 ATGTTATGGCAGCACAGTTAT ATGTTCTCCTGC ATCTGTATCTAGCACTacACaAGAAGTAT  
 11 3280 ATGTTATGGCAGCACAGTTAT ATGTTCTCCTGC ATCTGTATCTAGCACTgtACgAGAAGTAT  
 30 33 3306 ACATGtGGGTGGTCAGGTAAT tgTTTGTCTCTAC gTCTaTATCTAGCA ACCA AaTAT  
 16 3312 tCATGCGGGTGGTCAGGTAAT atTaTGTCTTACATCTGTgTTTAGCAGCA ACGA AgTAT  
 31 3251 gCATGCGGGTGGTCAGGTAATTG TTTtTCCTGaATCTGTaTTTAGCAG TGACGA AaTAT  
 35 18 3389 aCATtttGGGaaTaAtGTAATTGattgTaaTgactctatgtgcagTAcCAG TGACGAcacggTAT  
 con acaT---GGt-gt-agGTaATTg-at-tt-Tcctgcac-tct-t--c-AGcactgac-aagaagTAT  
 BE21-CGGTAT  
 40 6 3342 CCATTcCTGAA tCTACTACATACCCCCCGCACAGACC tCcACCCT tGTGTCCtCaaGC AC  
 11 3342 CCATTgCTGAA CCTACTACATACCCCCCGCACAGACCacCcCCCTacaGTGTCCgCctGC AC  
 33 3362 CCACTaCTGAAACTgcTgACATACA gACAGAC AAC gaTaacCGACC AC  
 16 3371 CCTCTcCTGAAAtTATTaggcagCA cttgGCC AACcACcCCgCCgCGACCcATAC  
 45 31 3310 CCT tTGCTggGATTGTtACaaAGCTACcaacaGCC AACaACACCcACCaTcGaaATTC  
 18 3454 CC gcTaCTcaGcTTGTtAAAcAGCTAC AgcACACCcCCtCacCGtATTC  
 50 con CCact-cTgaaa---ttgacatacAcccagcacagacc--c--caacAac-cctcc-Caacc-atac  
 CC---GCXACXCAGCXXG-BE21

6 3403 CA AGGAAGACGcAGTGcAaaACGCCCGCCTAGGAaAGC  
5 11 3406 CAC GGAAGACGGcGTGtCGGCGCCGCTAGGAaGCG  
33 3410 CACAAGC agcggcccAAACGacGACGAC cTgCAGacACCA  
16 3426 CA AAGCCgtCGCCTTGGGCACC GAAG AAACacaGACGAC TATcCAGCGACCAAGA  
10 31 3368 CA AAACCTtgCGCCTTGGGCACCaGtGAAGgtgTgCGGCggggGACGACGTCTActaAGCGACCAAGA  
18 3503 CA gcACCgtgtcCgTGGGCACC GcaAAGaccTAcGGC caGACGTC  
BE10-GGCGXGXCGGCGCCGCCXAG-BE10  
con CACAaaaccgtcgcccttgggcacc-g-gaaggcgtacgaagac-gacgacgtcc-cc-ag--accaaca  
15 6 3439 AGCACGaggaggtccaACaGTCCcCTtgCAACgCCTTGTTGTGGCCcACATTgGacCCGTGGACAGTG  
11 3442 AGCACG tggACCGTCCaCTaaCAACaCCcTGTGTGTGGCCaACATCaGatCCGTGGACAGTA  
20 33 3449 CAGACACCGCCcAgCCcCT tacaAAGcTGTtGTGCA gaCccCgCctTGGACaATA  
16 3481 tCAGAGCCAGACACCG GAAACCCCTgCCACaCCaActAAGTTGTGGACaGaGACTCaGTGGACAGTG  
31 3435 aCAGAGCCAGAGCaC aGAAACaCCcACcCCaAcAAGTTGTGGCAGgCGGACTCcGTGGACAGTG  
25 18 3548 GggtgctAcACGaCCTggACaCtgtggAcTcGcgGaGAaGcagCattGTGGAC cTG  
con acaaagccaga--ccgc-aaaCccct-c-acaccatgt-tttggtgcacagcggtccgTGGACagTg  
30 6 3507 gAAACcACAACcTCaTCACTaAC AATcACgACcAGCACCAAA GacGG AACACaG  
11 3504 cAATCaACAACATCGTCACTgAC AATtACAACAAGCACCAAA GAaGG AACAACTG  
33 3506 gAACAgcaCgtACTGCAACTAAcTGC aCAAAACAAGCAgCGGA cTgtGTGT AgTTc  
16 3548 cTCCAAcTcTcACTGCattTAACAGCT CACACAAAggACGGA tTaaCTGT AaTag  
35 31 3502 TCAACTgtggggTTaTCaGTGCAGCT gcatgCACAAAccAAACAA GGgCTGTCaGtTGTcc  
18 3604 TCAAC ccacTTcTcGTGCAGCTacacctacaggcaacAACAAaagacGGaaacTcGTaGTgg  
con c--aac-c--cactgc-actaaCagcta-at-c-aacaagcacca-Aa-----gggtgtcaaca-t-g  
40 6 3562 TaACAGTtCAGCTACGCCTATAGTGCAAtTtCAAGGTGAaTCCAATTGTTAAAgTGTtTTAGATATA  
11 3559 TcACAGTGCAGCTACGCCTATAGTGCAAcTgCAAGGTGAaTCCAATTGTTTAAATGTTTtTAGATATA  
45 33 3561 TAACgtTGCA CCTATAGTGcATTAAaAGGTGAaTcAaATAGTtTTAAATGTTTAAAGATA  
16 3603 TAACACTACA CCcATAGTACATTAAaAGGTGATGcTAATAcTTTAAATGTTTAAAGATA  
31 3563 TgcaACTACA CCTATAATACAcTTAAaAGGTGATGCAATAtaTTAAATGTTTAAAGATA  
50 18 3668 TaacACTAC gCCTATAATACATTTAAaAGGTGACagAAAcAgTtTTAAATGTTTAcGgTA  
con TaacacTaCagctacgCctATAgT-CAttTaaAAGGTGAAttcaAatagtTTAAATGTTTaaGaTata  
JJ20-catttaaaagggtgaNtcNaatagtttaaaatgtttaagatat

6 3630 GgCTaAATGACAgAcAcAgACATTaTTTGATTTAaATCaTCAACGTGGCAcTGGGCCTCctCaaAG  
 11 3627 GaCTgAATGACAAATATAAACATTGTTTGAaTTAgcATCtTCAACGTGGCATTGGGCCTCacCtgAG  
 5  
 33 3621 CAGATTaAAACcTATAAAgAgTTGTATAgTTCtaTGTCaTCCAcTGGCATTGGACCaGtgAcAAc  
 16 3663 TAGATTtAAAAgCAtTgtacATTGTATAcTgCAGTGTcGtCTACATGGCATTGGACAgGacAtAAT  
 31 3623 TAGGcTGTcAAAAAtATAaacAATGTATgaAcAAGTGTcATCTACATGGCATTGGACAtGtacagAT  
 10  
 18 3728 cAGatTGcgAAAAcATAgcgAccacTATagAgAtaTaTCATCCAcTGGCATTGGACA ggtgc  
 con g-agattt-aaaaa-Ata-aca-ttgTaT-a-t-a-t-TCaTC-AC-TGGCAtTGGaC-tg-cc-sa-  
 gcagatt-JJ20  
 15  
 6 3698 GCACCACATAAA cATGCCATTGTAACTgTAACAT ATgatAGTGAGGAACAAAGgCAaCAGTTTT  
 11 3695 GCACCACATAAAA ATGCAATTGTAAcAtTAACAT ATAgcAGTGAGGAACAACGtCAgCAaTTTT  
 33 3688 aAAAtagTAAAA ATGgAATTGTAACTgTAACATtTGtaAcTGAACAGCAACAAC AaatgTTTT  
 20  
 16 3730 GtAAAAcATAAAA gTGCAATTGTtAcAcTtACATATGatAGTGAATgGCAAC GtGACcAaTTTT  
 31 3690 GGAAAAcATAAAA TGCTATTGTaACcTAAcATATataAGTAcATCACAA AGAGACgAtTTTT  
 18 3791 aGgcAAtgaAAAAAcAGaATacTgACTgTAACATAccatAGTgaAaCACAA AGAacaaAaTTTT  
 25  
 con g-a-aacatAAAasatGcaATTgTaACTgTaACATatgatagt-aa-agcAAcaag--aacaaTTTT  
 6 3762 TAgAtgtTGTAAAAATACCcCCtACCATTAGcCA cAAacTGGGAtTTTATGTCaTgCAccTATTGTA  
 11 3759 TAaAcAgTGTAAAAATACCACCcACCATTAGgCAT AAGgTGGGgTTTATGTCAAtTaCATTtATTGTA  
 30  
 33 3752 TAggTAccGTAAAAATACCACC tACTgTgcaAAT AAG TACTGGATTAT  
 16 3794 TgtcTcaaGTtAAAAATACCA AAaACTaTtaCAGT GTC TACTGGATTAT  
 31 3754 TAAATACTGTaAAAAATACC tAAcACagTatCAGT GTCaacaggatatatgactATTTA  
 35  
 18 3856 TAAATACTGT  
 con TaaatactGTaaaaataaccaccaaaca-tagcaat-aaggtcgg-tttatgt-actg-atttattgta  
 6 3829 AttgttatatatgtaaAtgtgTaaATATATGgTATtgGTGTAatacaActgTACaTGTATGGAAgTgG  
 40  
 11 3826 A cCATTACaccTgtATATATG TATAtGTGTA CATAACATACgTGTATGGAgGTaG  
 33 3801 G aCATTa TAaGTGTA CATcACAagCcaaTATG  
 16 3843 G tCt  
 45  
 31 3812 gCc  
 18 3866 tgcaattccagatagtgtacaaatattggtgggataCa  
 con a-----g--catta----t--atatgggtatatgtgta--cataacaaacatgtatggaagtcg  
 50  
 55

6 3897 TGCCTGTACAAATAGCTGCAGGAACAACcAgcACATTcATAcT GCCTGTTaTaATTGCAT  
 11 3881 TGCCTGTACAAATTGCTGCAGcAACAACTAcAACATTGATATT GCCTGTTGTTATTGCAT  
 33 3833 TGCTGC TaAtTgtATAtAACcATGATATTtgTTTTG taTTaTGTTTaTaTT  
 16 3847 atATGA caAAtcTtgatacTgcATccAcAaCATTAcTgGcgTGCTTTTTG CTTTGCTTT GTGTG  
 31 3815 TaATGATtgAActaAatattTcTAcagtaAgCATT gTGctaTGCTTTTTG CTTTGCTTTGTGTG  
 18 3904 TgAcaATgtAAAtacAtatgcTgTAgTaccAatATgttatCacTtaTTTTTtatTTTGCTTTGTGT  
 con tgac-atacaa-ttgctgc-tgaacaacca-cA--tt-ata-TgctttttggccTtt-cTtttgtgtt  
 021-CTGCAGGAACAACCAGCACATTTCATACT GCCTGTTATAATTGCAT  
 6 3957 TTGttGTATGTTtTgTTAGcATcATACTTATtgTATggATATCTGAgTTTaTtGTgTAcACATCTGTG  
 11 3941 TTGcaGTATGTaTTcTTAGtATtgTACTTATAaTATTAATATCTGAcTTTgTaGTaTATACATCTGTG  
 33 3886 gT TTTtaTGctTaTcCtTATTAtTACGTCCtTTAATAcTTTCcaTTTCTACcTATgCtTggTTG  
 16 3911 CT TTT GTGTGTCTgCcTATTAATACGTCCgCTgcTttTGTCTGTGTCTACATAcaCATCATTA  
 31 3880 CTactaTTT GTGTGTCT tgTcATACGTCCaCTtgTgcTGTCTGTGTGggtATAtgCAaCAcTA  
 18 3971  
 con -tgctgtttg-tgtgt-tgcat--ta-tacgtccatt-atattttct-tttctgtatatacatc--tg  
 021-TTGTGTATGTTTGTAGCATCATACTTATTGTATGGATATCTGAGTTTATTGTGTACACATCTGTG-021  
 6 4025 CTaGTACTAACACTgCTTTTATATTaCTaTTGTGGCTgctATTAAACAACCCcTT GCAATTtTTcc  
 11 4009 CTGGTACTAACACTTCTTTTATATTGCTTTTGTGGCTttTATTAAACAACCCCTT GCAATTcTTT  
 33 3950 CTGGT gTTGGTATTgcTgcTtTggGtgTTTGTGG gAtCtCCTTTaaaAATT TTTT  
 16 3974 aTAaT ATTGGTATT acTaTTgTGGaTAACaGCAgCCTCTgCgTTTaG  
 31 3943 cTAtT ATTaaTtgT gaTtTTaTGGgTtAttGCAaCCTCTcCaTTacG  
 18 3971  
 con ctagtac-tt-atTTTTTTtatatttgcttttTgtggcttttatgaa-aac-cc-ttc-caatTTTT--  
 021-CTAGTACTAACACTGCTTTTATATTACTATTGTGGCTGTATTAAACAACCCCTT GCAATTTTTTCC-021

6 4092 TACTAACTCTactTGTGTGtTAcTgTCCcGCaTtgTATATACACTacTAtATTGT taccacA  
 11 4076 TACTAACaCTgtgTGTGTGCTATTTTCCTGCcTTTATATACACataTACATTGTgcaaacgcAAcaA  
 5 33 4005 T tTGCTATTT GTTgtTTTATAT tTA CCAATGaTgTgtATtAATtt  
 16 4021 gTGTtTTaTTGTATATAtaTATTT GTT TATA TA CCATTATTTTAAaTAcATAC  
 31 3990 tTGTtT TTGTATATAT gtTgt GTT TATA TAtattCCATTATTTgTAATtCATAC  
 10 18 3971  
 con tactaacactgtatat-tgctatTTTTcgTtc-tttatataactata-tccattgtttttaat-cata-  
 021-TACTAACTCTACTTGTGTGTTACTGTCCCGCATTGTATATACACTACTATATTGT TACCACA-021  
 6 4154 cAgcaatGATGcTAACaTGTCAaTTtAAATGATGGaGAT ACcTGGcTGggtTtTGTGGTGTtTatgTG  
 11 4144 TAATGgTGATGtTAACcTGCTACTTAAATGATGGtGAT ACATGGtTGTtTcTGTGGTGTtTtTcTG  
 15 33 4051 TCATGCACAgcaTAtgacacaACAagAgTAATGTATAT ACATGtaTaTATGTtTGTATATAtgTG  
 16 4075 ACATGCAC gcTTTTTAAttacataATGTATATGTACATAaTgtaATTGTT ACATATA  
 20 31 4044 ACATGCA tctTTTTTAA GTCAAcAgTaaCTTTTTT AC  
 18 3971 AtgcatgtatgtgtgctGcCATgtcccgCTTTTgccAtctgtctgta  
 con -catgcacatg-taac-t-t-Aattaaataatggagatgtacatggttg-tTtt-tg-t-t-tatgtg  
 25 021-CAGCAATGATGCTAACATGTCAATTTAATGATGGAGAT ACCTGGCTGGGTtTGTGGTGTtTATGTG-021  
 6 4220 CcTTTAtTTGTAGggaTgtTgGGgTTaTtTt gaTgCAcTAtAGaGCTGTACaAGGggaTAAAc  
 11 4210 CATTtTgTTGTAGcTGTacTtGGATTGTTGTT acTaCATTAcAGgGCTGTACATGGTAcTgAAA  
 33 4117 CA catGGTGgTGTtTTAAcATTGTTGTT gTATTTT AgtTTTTTTTTTTTGTa  
 30 16 4132 atTgTTGTATACCaTaActtactaTtTtTctTtTTATTTTcaTatAtaATTTTTTTTTTGT  
 31 4081 TtGTGTATAC TgTTgtTTgTatTgCTattggTaTggTaTTgg  
 18 4018 tgtgtgcGTaTgcATgggtattgggtatttTgtgtatatTgtggTaataacGTcccctgccacagcaTtc  
 35 con catTTt-tt-gtg-a-t-ttag--tt-tt-tt-t-tt-ttt-ttt--a--g-t-t-tttttt-ttt-t-  
 021-CCTTTATTGTAGGGATGTtGGGGTTATTATT GATGCACTATAGAGCTGTACAAGGGGATAAAC-021  
 6 4283 AcACcAAATGTaagAAGTGTA CAAAC aCAActgTAaTGatGATTATGTaactATGcattATAcT  
 40 11 4273 AaACTAAATGTgctAAGTGTAATCAAAcGCAAtacTAcTGTgGATTATGTGTATATGtcacATggT  
 33 4171 ttACTAA TAAAT AccTTTATaTttagcaGTGTAT  
 16 4196 TTgTTTgtTTgTTTTTA  
 45 31 4124 tattggTATtggTaTaaTaaacTTTTTtAcTTtTTTTTA  
 18 4086 acagtaTATgtaTtTgTtTtTtTaTTgccCaTgTTAcTattgcatatacatgctatatattgtctttaca  
 con a-actaaatgtattAagtgtaatt-t--cc-t--tttT-atgttgattAagtgatatg---tataact  
 021-ACACCAAATGTAAGAAGTGTA CAAAC ACAACTGTAATGATGATTATGTAACATATGATTATACT-021

6 4348 acTgATGGtGATTAT aTaTAtATGAAtTAGAGTAAACCgTTTTTATAtttgtaacaGTCTAtGc  
 11 4341 gaTaATGGaGATTATG TgTACATGAActTAGAGTAAACC TTTTTATAcagtgtgtgGTGTAcGt  
 5 33 4206 tATTATG  
 16 4214 ataaactgTTATTA  
 31 4164 TTATTA  
 10 18 4154 gtaattgtataggtgttttatacagtgattgtacattgtatattttgttttataccttttaTgcTt  
 con g-taatggagattatg---tatacatgaa-tagagtaaaacc-ttttttatatt-ttaat-gt-tatt-  
 021-ACTGATGGTgATTAT ATATATATGAATTAGAGTAAACCgTTTTTATATTGTAACAGTGTATGC-021  
 15 6 4413 TttgTATAccATggcacAtagTAGGGCCcCGacGACGcAAgCGTGCGTCAGCtACACAgCTATATCAAA  
 11 4405 TagtTATA tATAatgAAACcTAGGGCCAGCcaGACGtAAaCGTGCGTCAGCcACACAACTATATCAAA  
 33 4213 AGACACAAACGaTCTaCAAGgCGCA AgCGTGcATCtGCaACACAACTATAcCAAA  
 20 16 4228 CtttaacaATGCGACACAAACGtTCTgCAAAACGCACaAAACGTGCATCgGCTACcCAACTtTATAAAA  
 31 4170 C cATGCGgtcCAAACGcTCTaCAAAACGCACtAAACGTGCgTCTGCTACaCAATTaTATcAAA  
 18 4222 tttgtattTttGtaatAAaGtatggtAtccCaCcgTgccgcacgacgcaaacggggtTcggttaactg  
 25 con tttgtatat-aga-acaAacgt-c-gcaagacgc-gtaaactgtc-tc-gctacacaactatatcaaa  
 021-TTTGTATACCATGGCACATAGTAGGGCCCGACGACGCAAGCGTGCGTCAGCTACACAGCTATATCAAA-021  
 6 4481 CATGtAAactCACTGGAACATGcCCCCCAGATGTAATTCCCTAAgGTgGAgCacAacACcATTGCAGAT  
 11 4472 CATGCAAGGCCACTGGtACATGtCCCCCAGATGTAATTCCCTAAAGTtGAACAtAcTACTATTGCAGAT  
 30 33 4268 CATGCAAGGCCACAGGcAcCTGcCCACCCcGATGTTATTCCTAAAGTgGAAGGaAgTACcATaGCAGAT  
 16 4296 CATGCAAAcagGCAGGTACaTGTCCACCTgACaTTATACCTAAgGTtGAAGGcAaaAcTATTGCTGaa  
 31 4233 CATGtAAAgcAGCAGGTACTTGTCCATCaGACgTTATACCTAAaaTaGAACaTActAcCATTGCaGAC  
 35 18 4290 acTtatAtaaAaCAtGTAaacaatCtggtacatgTccACCTgAtgTtGttCcTaaggtggagGgcacC  
 con caTgcaAagccaCagGtAcatgtcCaccagatgttat-CCTaAagTtGaascataataccattGcagat  
 021-CATGTAAACTCACTGGAACATGCCCCCAGATGTAATTCCCTAAAGTGGAGCACAACACCATTGCAGAT-021



6 4549 CAAATATTAAATGGGGAAGtTTgGGGGTgTTTTTTGGaGGGTgGGTATaGGcACgGGttCcGGcAc  
 11 4540 CAAATATTAAATGGGGAAGcTTAGGGGTTTTTTTTGGTGGGTAGGTATTGGcACAGGggCTGGTAg  
 33 4336 CAAATtcttAAATATGGcAGTTTAGGGGTTTTTTTTGGTGGcTTAGGTATTGGcACAGGcTCTGGTtC  
 16 4364 CAAATATTAcAATATGGaAGTATGGGTGTaTTTTTTGGTGGGTAGGaATTGGaACAGGgTCgGGTAC  
 31 4301 CAAATATTaaggTATGGcAGTATGGGTGTtTTTTTTGGTGGGTgGGtATTGGgtCcGGCTCtGGTAC  
 18 4358 acgtTAGcAgataAaatattgcaatGgtcaagccTTGGTataTtttgggTGGacttGGCataGGTAC  
 con caaaTattaaaaatggaagttt-gGggtttttTTGGTgggTTaggtattGG-acaGGCctctGGtac  
 021-CAAATATTAAATGGGGAAGTTTGGGGGTGTTTTTTGGAGGGTGGGTATAGGCACGGGTTCGGGCAC-021  
 6 4617 TGGgGGTCGTaCtGGcTATgTtCCCTTacaAAcTCTgCaAAAcCTtCTATTACTaGtGGGCCtatgG  
 11 4608 TGGcGGTCGTgCaGGgTATaTACCCTTgGAAgcTCTCCcAAgCCTGCTATTACTgGgGGCCagCaG  
 33 4404 AGGtGGaAGgACTGGcTATgTACCTaTtGGtACTgacCCaCctACAGCTgCAAtccCctTGCagCCTa  
 16 4432 AGGcGGACGCACTGGgTATaTtCCatTgGGAaCAaGgCCTCCcACAGCTaCAGataCAcTTgctCCTg  
 31 4369 TGGgGGTCGCACTGGaTATgTcCCtcTtaGtACACGtCCTtCTACAGtaTCTgAggCAagTaTaCCTa  
 18 4426 TGGcaGTgGtACaGGgggTcgtaCagggtacAttCcattgggTgggsgtTCcaAaCAgtggTggaTg  
 con tGGcgGtcGtaCtGGgtaTgttcC-ttgggaAct-ctcc--ctacagctactaatacag-gcc-cctg  
 BE11-GAAGCXcXCCCAAGCCXGcXAXX-BE11  
 BE12-TATATACCCTTGGGAAGCXcXCCCAAGCCXGcXAX-BE12  
 021-TGGGGGTGCTACTGGCTATGTTCCTTACAAACTTCTGCAAAACCTTCTATTACTAGTGGGCCTATGG-021  
 6 4685 CtCGTCCtCCtGTGgTgGTGGAGCCTGTgGCCCCCTTCgGATCCaTCTATTGTGTCTtTTAATTGAaGAa  
 11 4676 CACGTCCgCCaGTGcTTGTGGAGCCTGTGCCCCCTTCcGATCCcTCCATTGTGTCTtTTAATTGAaGAa  
 33 4472 TACGTCCtCCgTtACTGTAGAcACTGTGGaCCTTtaGActCgTCTATAGTGTCaTTAATaGAAGAA  
 16 4500 TAAGACCcCCtTtAcAGTAGaTCCGTGgGcCCTTctGATCCtTCTATAGTtTCTTTAGTgGAAGAA  
 31 4437 TTAGACCcCCaGTTAgCaTtGACCTGTaGGtCCcTTGGACCCCTCTATAGTAagTcTTGTTGAAGAA  
 18 4494 TTgGtCCtACagTccCccGtggTtTaTtGaaCCTgTGGgCCCCaCagaccAtcTaTTGTTacAtta  
 con t-cGtCCtCagttac-gtaGagccTgTtGgcCCTt-gGa-cCctCtatagtgtcttTa-Ttgaagaa  
 BE26-CGXCCXCCGXXACXGAGAXa-BE26 JJ22-TCTATTGTGTCNTTAATNGAAGAA  
 BE27-GXCCXCCGXXACXGAGACACX-BE27 O22-GGATCCATCTATTGTGCTTTAATTGAAGAA  
 BE28-XCCXCCGXXACXGAGACACXGXXGGACCXXXAG-BE28  
 021-CTCGTCCtCCTGTGgTGGTGGAGCCTGTGGCCCCCTTCGGATCC-021

6 4753 TCgGCAATcATTAACGCAGGgGCC | TGAaaTtGTgCCCC | TgCACAcGGTGGgTTTAC  
 11 4744 TctGCTATTATTAATGctGGTGCACC | TGAggTgGTaCCCC | TACACAgGGTGGcTTTAC  
 5 33 4540 ACaAGTTTTATaGAgGCAGGTGCACCA | GccCCaTctATTCC | TACACcATCAGgtTTTga  
 16 4568 ACTAGTTTTATTGATGCTGGTGCACCAaCatCTGtaCCtTCCATTCCcCCagatgtATCAGgATTtag  
 31 4505 tCTGGaaTTgTTGATGTTGGTGC | ccCTGCTCCTAtaCCacacCCCTCTacaACATCTGGGTTTGA  
 10 18 4562 ataGaggactccagTCTgGtTaCatcaggTGCcCTAggCctacgttTaCTggcACgTCTGGGTTTGA  
 con -ctggtattatt-atGctGgtgCacca-ctgctgc-at--c---cccctcct-caccatctGGgTTT-a  
 TCTAGTNTTATTAATGCAGGTGCACC-JJ22  
 BE5-CAXXAACGCAGGGGCGCC-----XGAA-BE5  
 BE6-GGCAAXCAXXAACGCAGGGGCGG-BE6  
 15 BE7-GCAAXCAXXAACGCAGGGGCGCC-----XGAAAXXGXGCC-BE7  
 05-GTACCCCC TACACAGGGTGGCTTTAC  
 022-TCGGCAATCATTAAACGCAGGGGCGCC TGAaATTGTGCCCC TGCACACGGTGGGTTTAC-022  
 6 4812 aATtACATCctCTGAAaCaACTAcCCTGCAaTaTTgGATGT | aTCaGTT | ACTAgTCACACTA  
 11 4803 TATaACATCATCTGAAtCgACTACACCTGctATTTTaGATGT | gTCTGTT | ACcAaTCACACTA  
 20 33 4599 TgTTACTACATCTGCAGATACTACACCTGCaATTATTAATGTTtcaTCTGTTggggAgtcacTATTc  
 16 4636 TATTACTACTtCaaCtGATACcACACCTGC tATATTA | GATattaAtAaACTGTTA  
 25 31 4570 cATTgCTACaaCtGCaGAcCaACACCTGC aATTTTA | GAT | gtaACaAgTGTTA  
 18 4630 tATaaCatCtgCgGgtacaACTACACCTGCGgtTTTggatatcacaccttcgtctacctCtgtgtcTA  
 con taTtaC--CatCtgag--ACTACaCCTGCaatttTt-atgt--catctggt--tac-act--ta-Ta  
 05-TATAACATCATCTGAATCGACTACACCTGCTATTTTAGATGT GTCTGTT ACCAATCACACTA-05  
 30 022-AATTACATCCTCTGAaACAACTACCCCTGCAATATTGGATGT ATCAGTT ACTAGTCACACTA-022  
 6 4874 CtACTA GTaTaTTTtagAAATCCtgTcTTTACAGAACCtTCTGTAAcACAaCCCCAACCACCcGTG  
 11 4865 CcACTA GTgTgTTTcaAAATCCcTgTTTACAGAACCgTCTGTAAATACAgCCCCAACCACCTGTG  
 35 33 4667 aaACTATTtCTACACATtAAATCCCaCaTTTACTGAACCATCTGTAcTACaccCTCCAgCgCCTGCA  
 16 4692 ctACTgTTaCTACACATaTaAATCCCACTTTcACTGAcCCATCTGTATTGCAGCCTCCAACACCTGCA  
 31 4623 gCACACATgAaAATCCTACTTTTACTGATCCATCTGTATTGCAGCCTCctACACCTGCA  
 40 18 4698 tttccacaacCaatttTaccAATCCTgCaTTTtCTGATCCgTCcaTtaTtgAagtTCCacaAaCTGgg  
 con ctactatta-taca--TaaaAATCC-ac-TTtaCtGAaCCaTctgTaatacAgcctCcaccacCtGc-  
 05-CCACTA GTGTGTTTCAAAATCCCGTgTTTACAGAACCgTCTGTAAATACAGCCCCAACCACCTGTG-05  
 022-CTACTA GTATATTAGAAATCCTGTCTTTACAGAACCtTCTGTAAACACAACCCCAACCACCGTG-022  
 027-CTGCA

6 4939 GAGGCTAaTGGaCATATAtTaAtTCTGCaCCcACTgTAACgTCaCacCcTaTAGAgGaaATTCCttt  
 11 4930 GAGGCCAgTGGtCacATAcTtAtaTCTGCCCCaACaaTAACaTCCcCAACaTgTAGAAGACATTCCAcT  
 5 33 4735 GAAGCCtctGGaCATTtTATaTtTCTTCCcCTACTgTTAGcACACAAAgTTATGAAaACATaCCAAT  
 16 4760 GAAACtggAGGgCATtTTAcACTTTTCATCATCCACTATTAGtACACATAATTATGAAGAAAtCCtAT  
 31 4682 GAAACatCAGGTcATTtACTTTTCATCATCtCTATTAGcACACATAATTATGAGGAAATACCTAT  
 10 18 4766 GAggtggCAGGTaATgTAtTgTtggtTaCccCtaCateTgGaACACATgggTATGAGGAAATACCTtT  
 con GA-gc--c-GGtcAttTa-ta-TttcttC-cC-aCtattag-aCaCAtaattatGA-gAaAT-CCtat  
 05-GAGGCCAgTGGTCACATACTTATATCTGCCCCAACAAATACATCCCAACATGTAGAAGACATTCCACT-05  
 022-GAGGCTAATGGACATATATTAAATTTCTGCACCCACTGTAACTGCACACCCCTATAGAGGAAATTCCTTT-022  
 027-GAAGCCTCTGGACATTTTATATTTTCTTCCcCTACTGTTAGCACACAAAGTTATGAAaACATACCAAT-027  
 15 6 5007 AGAtACTTTTGTgGTATCaTCTAGTGATAGcGGtCCTACATCCAGTACcCCTgTTCCTgGTaCTgcaC  
 11 4998 AGAcACTTTTGTGTATCCTCTAGTGATAGTGGaCCTACATCCAGTACTcCCTcTTCCTcGTgCTtTtC  
 20 33 4803 GGATACcTTTGTGTtTCCACAgACagTAGTAatGTAACATCaAGCAGcCCCATTCCAGGGTCTCGCC  
 16 4828 GGATACATTTATTGTtagCACAAACccTAACaCaGTAACtAGTAGCACACCCATaCCAGGGTCTCGCC  
 31 4750 GGATACATTTATTGTtTCTACTAAtaTGAAaAcaTAACaAGTAGCACACCCATtCCAGGGGTGCGCC  
 18 4834 acAaACATTTGcTtCTTCTgGTAcggggGAggAacccAttAGTAGtACcCCATgCCTactGTGCGgC  
 25 con -gAtACaTTTgttgtttccactaatgata---aac-aAca--TAG-AC-CCcattCC-gg-gctcgcc  
 05-AGACACTTTTGTGTATCCTCTAGTGATAGTGGACCTACATCCAGTACTCCTCTTCCCTCGTGTCTTTT-05  
 022-AGATACTTTTGTGGTATCATCTAGTGATAGCGGTCTACATCCAGTACCCCTGTTCCTGGTACTGCAC-022  
 027-GGATACCTTTGTGTtTCCACAGACAGTAGTAATGTAACATCAAGCACGCCcATTCCAGGGTCTCGCC-027  
 30 6 5075 CTCGGCCTCGtGTGGGccTaTATAGTCGTGCaTTgCaCAGGTgCAGGTACaGACCcTGCaTTtTt  
 11 5066 CTCGGCCTCGgGTGGGTTTgTATAGTCGTGccTTaCagCAGGTACAGGTACgGACCCcGCgTTTtTg  
 33 4871 CTGTGGCAGCCcTGGTTTATATAGTCG CAAtACcCAACAGGTTA AGGTTGTtGACCCTGC  
 16 4896 CaGTGGCAGCCCTAGGaTTATATAGTCG CACaACACAACAGGTTA AAGTTGTaGACCCTGC  
 35 31 4818 GTcctGCACGtTtTAGGgTTATATAGT AaGGCtACACAACAAGTaA AAGTTaTTGatCCaac  
 18 4902 GTgtaGCAGGTccccGccTtTACAGT AgGGCctacCAACAAGT gtcAGTggcTaAcCCtga  
 40 con ct-tggCacGtct-gG-tTaTAtAGTcgtgc-atg---a--caaCagGTtaca-gttgttga-cctgc  
 05-CTCGGCCTCGGGTGGGTTTGTATAGTCGTGCCCTTACAGCAGGTACAGGTACGGACCCCGCGTTTtTG-05  
 022-CTCGGCCTCGTGTGGGCCTATATAGTCGTGCATTGCACCAGGTGCAGGTACAGACCCCTGCATTtCTT-022  
 027-CTGTGGCACGCCCTGGTTTATATAGTCG CAATACCCCAACAGGTTA AGGTTGTtGACCCTGC-027  
 45  
 50  
 55

6 5143 TCCACtCCtCAaCGcTTaattACaTAT GATAACCCCTGTaTATGAA GGgGAgGATG  
 11 5134 TCCACgCCCaCAGCGaTTggTAACTTAT GACAACCCCTGTcTATGAA GGaGAaGATG  
 5  
 33 4932 TTTTtTAACatCgCCTCaTAAACTTATAACATATGATAATCCTGCATTtGAAAgtTtGAccttGAaG  
 16 4957 TTTTgTAACCaCTCCcAcTAAACTTATTACATATGATAATCCTGCATATGAAGGTaTAGATgtgGata  
 31 4879 GTTTCtTAgtgCTCCAAaAcAgCTAATTACATATGaaaACCCCTGCcTATGAAacTgTAAATGctGAaG  
 10  
 18 4963 GTTTCtTAacgTCCAtcctcctTAATTACATATGACAACCC gGCctttG  
 con ttttct-accactcctta-taacttATTacatatGATAACCCtgcatatgaaagt-taga-gc-gatg  
 05-TCCACGCCACAGCGATTGGTAACTTAT GACAACCCCTGTCTATGAA GGAGAAGATG-05  
 022-TCCACTCCTCAACGCTTAATTACATAT GATAACCCCTGTATATGAA GGGGAGGATG-022  
 027-TTTTtTAACATCGCCTCATAAACTTATAACATATGATAATCCTGCATTtGAAAgtTTGACCTGAAG-027  
 15  
 6 5198 TtAGTgTACAATTtAgTCATGAtTCTA TaCACAATGCACCTGATGAgGctTTTATGGAcATa  
 11 5189 TaAGTTTACAATTtAccCATGAgTCTA TcCACAATGCACCTGATGAaGcATTATGGATATT  
 33 5000 AcACATTACAATTtcaaCATAgTGATA TatcaccTGCTCCTGATCCTGACTTtTaGATATT  
 20  
 16 5025 AtACATTATAtTTTtCtagTAAATGATAatagtaTTAATATAGCTCCaGATCCTGACTTtTgGATATA  
 31 4947 AatCTTTTACTTTTC caATAcacCgCaTAATATAGCcCCTGATCCcGACTTtTaGATATT  
 18 5013 AgcCTgTggACacTaCattaacattTgacCtCgTAgTgatGttCCTGATtCaGATTTTaTgGATATT  
 25  
 con a-acttTAcAattTac-catsattaTaatt-ctcttaataatGctCCTGATcc-GacTTTaTgGATAtT  
 05-TAAGTTTACAATTtACCCATGAGTCTA TCCACAATGCACCTGATGAAGcATTtATGGATATT-05  
 022-TTAGTGtACAATTtTAGTCATGATTCTA TACACAATGCACCTGATGAGGCTTTTATGGACATA-022  
 027-ACACATTACAATTtCAACATAGTGATA TATCACCTGCCTGATCCTGACTTtCTAGATATT-027  
 30  
 6 5260 ATTcGcttTgCAcAGACcTGCCaTTgCGTCCcGACGtGGcCTTGTGCGGTacAGTCGCATTGGaCAACG  
 11 5251 ATTaGAcTACATAGACCaGCTATaACGTCCaGACGgGGTCTTGTGCGTTTtAGTCGCATTGGgCAACG  
 33 5062 ATTGCATTACATAGGCCtGCTATTaCaTCTcGtaGacaTAcTGTGCGTTTtAGTAGAgTaGGTCAAAA  
 16 5093 gTTGcttTACATAGGCCaGcATaACCTCTaGgcGtAcTggcaTTAGgTAcAGTAGAaTtGGTAATAA  
 35  
 31 5009 ATaGcATTACATAGGCCTGCCcTtACCTCacGtaGgAacACTGTtAGaTATAGTAGAcTAGGTAATAA  
 18 5081 ATccgtcTACATAGGCCTGcctTaACaTCCaGgcGtgggACTGTTcGcTtTAGTAGAtTAGGTcAacg  
 40  
 con aTtg--tTaCaTAgGcCtGCTaT-aC-TCC-G-cGtggtactgT-cG-T-tAGTaGaaT-GGtcAa--  
 JJ24-TACATAGGCCCTGCTATAACNTCCAGNCGTGGTNNtGTGCGNTTTAGTAGA-JJ24  
 05-ATTAGACTACATAGACCAGCTATAACGTCCAGACGGGGTCTTGTGCGTTTtAGTCGCATTGGGCAACG-05  
 016-ACTGTGCGTTTtAGTAGAGTAGGTCAAAA-016  
 022-ATTGCTTTGACAGACCTGCCATTGCGTCCCGACGTGGCCTTGTGCGGTACAGTCGCATTGGACAACG-022  
 027-ATTGcATTACATAGGCCTGCTATTACATCTCGTAGACATACTGTGCGTTTtAGTAGAGTAGGTCAAAA-027  
 028-GTAGACATACTGTGCGTTTtAGTAGAGTAGGTCAAAA  
 45  
 50  
 55

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55



6 5713 GGTCTGgaTTtTATTTCATCCTgCATGGTatTTTGCACGtAaACGCCGTAAACGTATTCCCTTATT  
 11 5695 GGTCTGACTTcTATTTCATCCTACATGGTAcTTTGCACGcAGACGCCGTAAACGTATTCCCTTATT  
 33 5518 GGTGCTGACTTTgTTTACATCCTAGTTATTtTATTACGtcGcaGgCGTAAACGTTTTCCATATTT  
 16 5561 GcaGGTGACTTTTATTTACATCCTAGTTATTacATGTTACGaaaCGACGTAAACGTTTACCATATTT  
 31 5477 GGgGGTGATTTTATTTCACcCTAGTTATTATATGTTAaaAcgtCGACGTAAACGTGTatCATATTT  
 18 5537 GGtacacATTatTATTGtggCCattaTATTAtTtaTtcctaagaaACGTAAACGTGTtcCctATTT  
 con GgtgctgacTtttaTTTgcatCCtag-TatTat-Ttttacgta-acgaCGTAAACGT-TtcC-TatTT  
 JJ25-CGTAAACGTNTTCCCTATTT  
 PCR2-CGTTTTCCATATTT  
 027-GGTGCTGACTTTGTTTTACATCCTAGTTATTTTATTTTACGTCGCAGGCGTAAACGTTTTCCATATTT-027  
 6 5781 TTTTtCAGATGT GCGCGCCTAGCGACAGCAGATATATGTGCCTCCTCtAACCTGTATCC  
 11 5763 TTTTACAGATGT GCGCGCCTAGCGACAGCAGATATATGTGCCTCCTCCcAACCTGTATCC  
 33 5586 TTTTACAGATGTCcgTgTGGCGGCCTAGTGAGGCCACAGTgTACcTGCCCTCCT GTaCCTGTATCT  
 16 5629 TTTTtCAGATGTCTCTtTGgctGCCTAGTGAGGCCACTGTCTACTTGCCCTCCT GTCCcAGTATCT  
 31 5545 TTTTtCAGATGTCTCTGTGGCGGCCTAGcGAGGCTACTGTCTACTTaccACCT GTCCcAGTGTCT  
 18 5605 TTTTgCAGATGgCTtTGTTGGCGGCCTAGtGAcAaTAcCGTaTatcTtCCACCT cctcCtGTGgCa  
 con TTTTACAGATGtctctgtgTGGGgCCCTAG-GA--ccACaGTaTA--TgCCTCCTcc-gtccCtGTatCt  
 TTTTNCAGATGTCTNTGTGGCGGCCTAGTGA-JJ25  
 PCR1-CAGATGTCTCTGTGGCGGCCTAGTG-PCR1  
 TTTTGCAGATG-PCR2  
 027-TTTTACAGATGTCCGTGTGGCGGCCTAGTGAGGCCACAGTGTACCTGCCTCCT GTACCTGTATCT-027  
 6 5843 AAaGTTGTTGCCACGGATGcTATGTTActCGCACCAACATATTTTATCATGCCAGCAGTTCTAGACT  
 11 5825 AAgTTGTTGCCACGGATGcTATGTTAaACGCACCAACATATTTTATCATGCCAGCAGTTCTAGACT  
 33 5651 AAaGTTGtCAGCACTGATGAATATGTgtCtCGCACAAgCATtTATTATtATGcTGGtAGTCCAGACT  
 16 5694 AAgTTGTAAGCACGGATGAATATGTTgCACGCACAAACATATATTATCATGCAGGaAcaTCCAGACT  
 31 5610 AAAGTTGTAAGCACGGATGAATATGTaACACGaACCcAACATATATTATCAcGCAGGCAGtgCTAGgCT  
 18 5670 AgAGTTGTAAatACcGATGAtTATGTgACtCcAcAaAgCATAtTTATCATGcTGGCAGctCTAGatT  
 con AAaGTTGTaagcACgGATGaATATGTTac-CgcAC-AaCATaT-TTATcAtGC-gGcAgttCtAGact  
 JJ26-GTTGTNANCACGGATGANTATGTTACTCGCACAA-JJ26  
 PCR3-AAAGTTGTAAGCACGGATGAATATGT-PCR3  
 LCR1A-AAAGTTGTAAGCACGGATGAATATGT-LCR1A  
 LCR1B-TGCACGCACAAACATATATTATCA-LCRB  
 LCR1B'-ACGTGCGTGTGTTGTATATAATAGTA-LCRB'  
 027-AAAGTTGTCAGCACTGATGAATATGTCTCGCACAAAGCATTATTATTATGCTGGTAGTTCAGACT-027  
 6 5911 tCTTGCaGTGGGACATCctTATTtTCCaTaaaa cggGcTAA C AAAA CtGTTGTgC

11 5893 CCTTGCTGTGGGACATCCATATTactCTATcAAAAA aGtTAA C AAAA CAGTTGTAC  
 5 33 5719 tCTTGCTGTGGGcCATCCATATTTTCTATTAAAAATCCTAcTAA CgctAAAAAAATATTgGTAC  
 16 5762 aCTTGcAGTTGGaCATCCcTATTTTcCTATTAAAAACCTAACAAT AAcAAAA TATTaGttC  
 31 5678 gCTTACAGTaGGCcCATCCATATTAAtCcATacctAAAtCTGACAATccTAAaAAAA TAGTTGTaC  
 10 18 5738 atTaAcTGTtGGtaATCCATATT TtagggttcCTGcaggTggTggcAAaAagcagGaTaTtC  
 con -cTtgC-GTtGGAcATCCaTATTtttctaTtaaaaaaacctgcta--caacaaaAaa-tagttgTaC  
 JJ27-GTTGGACATCCATATTT-JJ27  
 027-TCCTGCTGTGGCCATCCATATTTTCTATTAAAAATCCTACTAA CGCTAAAAAATTATTGGTAC-027  
 15 6 5967 CAAAGGTGTcAGGATATCAATAcAGgGTaTTTAAGGTgGTGTTaCCAGATCCTAACAAaTTTGCATTg  
 11 5949 CAAAGGTGTcTGGATATCAATATAGaGTgTTTAAGGTaGTGTTgCCAGATCCTAACAAGTTTGCATTa  
 20 33 5784 CcAAAGTATCAGGcTTgCAATATAGGGTtTTTAGGGTcCgTTTACCAGATCCTAATAAaTTTGGATTt  
 16 5824 CtAAAGTATCAGGATTACAAATAcAGGGTATTTAGaaTaCaTTTACcTGACCCcAAATAAgTTTGGtTTT  
 31 5743 CaAAGGTgTCAGGATTACAAATATAGGGTATTTAGGGTtCgTTTACCAGATCCAAAcAAATTGGaTTT  
 25 18 5800 CtAAGGTtTcTgCaTAcCAATATAGaGTATTAGGGTgCagTTACcTGACCCAAaTAAATTGGtTTa  
 con CaAAGGTgTCaGgaT--CAATAtAGgGTaTTTAaggT-c--TTaCCAGATCCtAA-AAaTTTGGaTTt  
 JJ28-CAATATAGGGTATTTAGGGTNCNGTTACC-28 30-AAATAATTGGATTN  
 PCR4-GTTATATCCCATAAATCCCATGTTAA-PCR4PCR5-TTATTTAACCAAAAA  
 027-CCAAAGTATCAGGCTTGCAATATAGGGTTTTTAGGGTCCGTTTACCAGATCCTAATAAATTGGATT-027  
 30 6 6035 CCTGAcTCgTcTcTtTtGatCCcACaAcCGTTTAGTATGGGCcTGcACAGGccTaGAGGTgGG  
 11 6017 CCTGAtTCaTCCCTgTTTGACCCcACTACACAgCGTTTAGTATGGGCgTGcACAGGgtTgGAGGTAGG  
 35 33 5852 CCTGACACCTCCTTTTATAACcCtGATACACaAcGaTTAGTATGGGCcTGtGTAGGccTTGAaATAGG  
 16 5892 CCTGACACCTCaTTTATAATCCaGATACACAgCGgcTgGTTTGGGCCTGTGTAGGTgTTGAGGTAGG  
 31 5811 CCTGATAcATCTTTTATAATCCTGAaAcCAACGcTTAGTTTGGGCCTGTGTGGTtTaGAGGTAGG  
 40 18 5868 CCTGATAcTagTaTTTATAATCCTGAaACaCAACGtTTAGTgTGGCCTGTGcTGAGTgGaaTtGG  
 con CCTGA-aC-tcttTtTataAtCCtga-ACaCAaCGtTtAGTaTGGGCcTGtg-aGG--T-GAggTaGG  
 CCTGACACCTCNNTTATAAT-JJ30  
 GGACTGTGGA-PCR5  
 027-CCTGACACCTCCTTTTATAACcCtGATACACaACGATTAGTATGGGCATGTGTAGGCCTTGAAATAGG-027  
 45 50 55





6 6304 tAATACacCTGTACaggeTGGTGACTGCCCGCCcTTaGAACCTTATTACCAGTGTTATACAGGATGGcG  
 11 6289 aAATACctCTGTACaAaaTGGTGACTGCCcCCgTTgGAACTTATTACCAGTGTTATACAGGATGGgG  
 33 6124 TAaTGCAGcACcTgCcaATGATTGTCCACCcTTAGAACTTATAAAaTACTaTTATTgAGGATGGTG  
 16 6164 cAATgtTGCAGTAAaTCCaGGTGATTGTCCACCATTAGAgTTAATAAAcACAGTTATTcAGGATGGTG  
 31 6083 tAAcaatGCTaTtACCcCtGGTGATTGTCCcCCATTAGAAATTAAAAAaTtCAGTTATaCAAGATGGgG  
 18 6140 atcgegTcCTtTatCaCagGGcGATTGGcCCcCTTTAGAAcTtAAAAAcaCAGTTtTggAAGATGGcG  
 con -aata-tgCtgtac--cctgggTgATG-CC-CC-TTaGAacTtAtaAaac-gTTaTacAgGATGGtG  
 JJ36-GATGGTG  
 027- TAATGCAGCACCTGCCAATGATTGTCCACCTTTAGAACTTATAAAaTACTATTATTGAGGATGGTG-027  
 6 6372 AtATGGTTGACACAGGCTTTGGTGCTATGAATTTGCTGAtTTgCAgACCAATAAATCaGATGTTCCt  
 11 6357 AcATGGTTGATACAGGCTTTGGTGCTATGAATTTGCAGAcTTaCAaACCAATAAATCgGATGTTCCc  
 33 6189 ATATGGTgGACACAGGAtTTGGTtgcATGGATTTTaaAACATTgCAGGCTAAATAAAAGTGATGTTCCt  
 16 6232 ATATGGTtCATACTGGCTTTGGTGCTATGGAcTTTACTACATTACAGGCTAAaCAAAAGTGAAgTTCCa  
 31 6151 ATATGGTTGATACAGGCTTTGGGAGCTATGGATTTTACTgCtTTACaAGAcACTAAAAGTaAGTTCCT  
 18 6208 ATATGGTaGATACTGGaTatGGtGcCATGGAcTTTAgTaCaTTgCAAGAtACTAAATGTgAgGTaCCa  
 con AtATGGTtGAtACaGGcTtTGGTgctATGgAtTTTact-caTT-CA-gc-AaTAAA-gtgAtGTTcCT  
 ATATGGTTGATACAGGCTTTGGTGCTATGGA-36 37-CATTNCANGCNAATAAANGTGATGTTCCCT  
 027-ATATGGTGGACACAGGATTtTGGTtGCATGGATTTTAAACATTGCAGGCTAAATAAAAGTGATGTTCCCT-027  
 6 6440 aTTGAcATaTGTGGcACTaCaTGtAAATATCCaGATTATTTaCAAAATGGCTGCAGACCCaTATGGTGA  
 11 6425 cTTGATATTTGTGGaACTgtcTGCAAAATATCCtGATTATTTgCAAAATGGCTGCAGACCCCTTATGGTGA  
 33 6257 aTTGATATTTGTGGcAgTAcATGCCAAATATCCAGATTATTTaAAAAATGAcTagtGAgCCTTATGGTGA  
 16 6300 cTGGATATTTGTAcATCTATTGTCAAAATATCCAGATTATaTAAAAATGGTgtCaGAaCCATATGGCGA  
 31 6219 TTGGAcATTGTAAaTCTATTGTAAATATCCAGATTATcTTAAAAATGGTTGCTGAgCCATATGGCGA  
 18 6276 TTGGAtATTGTcAgTCTATTGTAAATATCCtGATTATtTAcAAATGtcTGCAGAtCCTTATGGgGA  
 con -T-GATAtTtGTgg--GTattTG-AAATATCCaGATTATtTa-AAATGgctgcaGA-CC-TATGGtGA  
 NTNGATATTTGT-JJ37 JJ38-AAATATCCAGATTATTTANAAATGG-JJ38  
 027-ATTGATATTTGTGGCAGTACATGCCAAATATCCAGATTATTTAAAAATGACTAGTGAGCCTTATGGTGA-027

5

10

15

20

25

30

35

40

45

50

55



6 6780 ATACCACACGCAGTACcAAcATGACaTtATG TGCATCcGTaaCTACATCTtCcACATACAcC  
 5 11 6765 ATACCACACGCAGTACaAAATATGACAcTATG TGCATCTtGTgtCTAAATCTgtTACATACAcT  
 33 6597 ATACCACtCGCAGTACtAAATATGACtTTATG caCAcaAGTaaCTAGtGAcAgTACATATAAA  
 16 6640 ATACTACACGCAGTACaAAATATGTCaTTATGTGCT GccATAtCtCtACTtcaGAaACTACATATAAA  
 10 31 6559 ATACCACACGtAGTACCAATATGTctgtTtGTGCT GCAATtgCaAacagtGATACTACATtTAAA  
 18 6616 ATACCACtCccAGTACCAAtTtTaaCaaTaTGTGCTtctaCacagtCtccgtgtaccTgggcaATaTgAt  
 con ATACCcACaGgcAGTACcAAaTgaCatTaTGTgtct--tgCa---g-aacta-ag-tactacATataaa  
 ATACC-JJ41 C16-CATCCGTAACATACATCTTCCA-C16  
 15 ATACCACACGCAGTAC-C1 C17-TCTGTGTCTAAATCTGTCTACA-C17  
 ATACTACACGCAGTAC-C7 C20-CACACAAGTAACAGTACAG-C20  
 ATACCACtCGCAGTAC-C8 C23-CAGTCTCCTGTACCTGGG-C23  
 ATACTACTCGCAGCAC-C10 C31-TTGCAACAGTGATACTACATT-C31  
 ATACTACCCGTAGTAC-C11  
 ATACTACCAGAAGCAC-C12  
 ATACTACTAGAAGCAC-C13  
 20 ATACCACACGTAGTAC-C14  
 ACCTACCCGCAGTAC-C15  
 027-ATACCACtCGCAGTACTAATATGACTTTATG CACACAAGTAACAGTACAGTACATATAAA-027  
  
 6 6842 AATTCTcGATTATAAaGAgTACATGCGtCATGTGGAaGAGTaTGATTTTACAaTTTATTTTCAATTaTG  
 25 11 6827 AATTcAGATTATAaGGAATACATGCGcCATGTGGAgGAGTtTGATTTTACAGTTTATTTTCAATTgTG  
 33 6659 AATgaAAATTTTAAaGAATaTATAaGACATGTtGAaGAATATGATcTACAGTTTgTTTTTCAACTaTG  
 16 6705 AATAcTAAcTTTAAgGAGTAccTAcGACATGGgGAGGAATATGATTTACAGTTTATTTTCAACTgTG  
 30 31 6624 AgTAgtAAcTTTAAaGAGTATtTAAGACATGGTgAGGAATtTGATTTACAaTTTATaTTTCAGTTaTG  
 18 6684 gCTAccAAaTTTAAgCAGTATagcAGACATGtTGAGCAATaTGATTTgCagTTTATtTTTCAGTTgTG  
 con aaTactaAtTtTAA-gAgTA-ata-GaCATGt-GAgGAaTaTGATtTaCagTTTaTtTTTCAaT-TG  
 027-AATGAAATTTTAAAGAATATATAAGACATGTTGAAGAATATGATCTACAGTTTGTtTTTCAACTATG-027  
 35  
 6 6910 TAGCATTACATTgTCTGCTGAAGTaATGGCCTATATtCACACAATGAATCCcTCTGTTTGGGAaGACT  
 11 6895 TAGCATTACATTATCTGCAGAAgTcATGGCCTATATaCACACAATGAATCCcTCTGTTTGGAgGACT  
 40 33 6727 CAAAgTTACCTTAACtGCAGAAgTTATGACATATATtCATgCTATGAATCCagaTATTTTaGAaGAtT  
 16 6773 CAAAATAACCTTAACtGCAGACGTTATGACATAcATaCATtCTATGAATtCcaCTATTTGGAgGAcT  
 31 6692 CAAAATAACaTtATcTGCAGACaTaATGACATATATTCaAGTATGAATcCtgCTATTTTGGGAaGATT  
 45 18 6752 tActATtAcTtTAaCTGCAGAtgTtATGtCcTATATTCAtAGTATGAATagcagTATTTTaGAgGATT  
 con -AaaaTtACaTTa-CTGCaGAagTtATGAc-TAtAtTCA-actATGAATccc-cTaTTTgGA-GA-T  
 027-CAAAGTTACCTTAACtGCAG-027  
 50  
 55

6 6978 GGAAC TTTGGg TTATCGCCTCCcCAAATGGTACAtTaGaaGATACcTATAGgTATGTgCAGTCACAG  
 11 6963 GGAAC TTTGG TTTATCGCCTCCcCAAATGGTACAcTgGAGGATACtTATAGaTATGTaCAGTCACAG  
 33 6795 GGcAa TTTGG TTTAaCACCTCCtCCAtcTGcTAgttTAcAGGATACcTATAGGTTTGTtaccTctCAG  
 16 6841 GGAAT TTTGG TcTAcAACCTCCcCCAggAGGcACacTAGAaGATACtTATAGGTTTGT aaCCcCAG  
 31 6760 GGAAT TTTGG GatTgaCCaCaCctCCctCAGGtC TTTGGAGGATACcTATAGGTTTGTcaccTCaCAG  
 18 6820 GGAAc TTTGG TgtTcCCcCCcCCcCaaCtacTagTTTGGtGGATACaTATcGtTTTGTacaaTctgtt  
 con GGAAc TTTGG TtTa-c-cCtCCcCCaactggtac-tT-gagGATACcTATaGgTtTGTaca-tCacag  
 C21-TTGT AACCCAG  
 C34-TTTGT AACCCAG  
 C35-GTTTGT AACCCAG

6 7046 GCCATTACCTGTCAaAAgCCCACtCCTGAAAAGGAAAAGCcaGA TCCCTATAAGaAccTtAGTT  
 11 7031 GCCATTACCTGTCAgAAACCCACACCTGAAAAAGAAAAaCagGA TCCCTATAAGgAtaTgAGTT  
 33 6863 GcTATTACgTGTCAAAAAaCagtACCTCCAAAGGAAAAGGAAGA CCCCTTAggtAAATAtACaT  
 16 6906 GCaATTgCtTGTCAAAAAaCaTaCACCTCCAgcaCCtAAaGAAGATgatCCCCTTA AAAAATACAcTt  
 31 6828 GCcATTACaTGTCAAAAAaCTGcCCcCaAaagCCcAAgGAAGAT CCaTTTA AAgAtTAtgtaT  
 18 6888 GCtATTACcTGTCAAAAaggaTGctgCaCcggtgaaAAaTaAgGAT CCcTaT gAtaAgTtaaagT

con GCcATTaCcTGTCAaAAacct-cacCtc-aaaggaaAaggaaGat---cCcttaaa-aA-t-ta-tT  
 GCAATTGCT-C21 C18-CATACACCTCCAGCACCTAA-C18 JJ46-T  
 C19-GGATGCTGCACCGGCTGA-C19  
 C22-AAAAACAGTACCTCCAAAGGA-C22  
 C27-TTTTGTCTATGGAGGTTTCCT-C27  
 C24-CACACCTGAAAAAGAAAAACAG-C24  
 C28-GTGTGGACTTTTCTTTTGTGTC-C28  
 C25-CTCCTGAAAAGGAAAAGCCA-C25  
 C26-GAGGACTTTTCCTTTTCGGT-C26  
 C29-CCAAAAGCCCAAGGAAGAT C-C29  
 C30-CAAAAAGCCCAAGGAAGAT C-C30  
 C32-CAGAAACCCACACCTGAAAAAGA-C32  
 C33-AGAAACCCACACCTGAAAAAGAA-C33  
 GCAATTGCT-C34 O23-GGA TCCCTATAAGGATATGAGTT  
 GCAATTGCT-C35 O15-GGAT CCCTAT GATAAGTTAAAGT

6 7110 TTTGGGAGGTTAACTTAAAGAAAAAGTTTCTAGTGAATTgGATCAGTaTCCcttTgGGACGcAAGTTT  
 11 7095 TTTGGGAGGTTAACTTAAAGAAAAAGTTTCAAGTGAATTAGATCAGTTTCCccttGGACGtAAAGTTT  
 33 6927 TTTGGGAAGTgGATTAAAGGAAAAATTTTCAGCAGAtTTAGATCAGTTTCCCTTgGGACGCAAGTTT  
 16 6973 TTTGGGAAGTAAATTTAAAGGAAAAAGTTTCTGCAGAcceTAGATCAGTTTCCCTTaGGACGCAAAATTT  
 31 6892 TTTGGGAgGTtAAATTTAAAGAAAAAGTTTCTGCAGAtTTAGATCAGTTTCCaCTgGGtCGCAAAATTT  
 18 6952 TTTGGAACTGgATTAAAGGAAAAAGTTTCTtAGAcTTAGATCAATaTCCcCTtGGAcGTAAATTT  
 con TTTGGGAgGTtAAATTTAAAGAAAAAGTTTCTgcAGa-tTAGATCAGTtTCCt-TgGGAcGcAA-TTT  
 TTTGGGAGGTTAACTTAAAGAAAAAGTTTCTGCAGANTTAGATCA-JJ46  
 C2-GATCAGTTTCCYYTKGGACG-C2  
 C3-GATCAGTWTCCYYTKGGACG-C3  
 C7-CTAGTCAGGRRAMCCTGC-C7  
 015-TTTGGAATGTGGATTAAAGGAAAAAGTTTCTTTAGACTTAGATCAATATCCCTTGGACGTAAATTT-015  
 023-TTTGGGAGGTTAACTTAAAGAAAAAGTTTCAAGTGAATTAGATCAGTTTCCCTTGGACGTAAAGTTT-023  
 6 7178 TT gTT aCAAAGTGGATATaGgGGACGGtCcT  
 11 7163 TTA TT gCAAAGTGGATATcGAGGACGGaCgT  
 33 6995 TTA TTAcagGcAggtctttaaagcAaaAcctaaacttaaACGtgCAGcccccaCAtCcACCCgCA  
 16 7041 TTA cTACAAGCAGGATtgAaCGCcAaACCAAAATTTAcAttAGGaAAACGaaaagCTaCACCcAcCA  
 31 6960 TTA tTACAGGCAGGATatAGGGCacgtCCtAAATTTAAAgCAGGtAAACG TagTgCACCC t  
 18 7020 TT ggTtCAGGCTGGATtgcGtcgcaagCCcAccaTaggcccCtcGcAAACG T tctg  
 con TTA--ta--agcaggattgagggcAAAAacAAAAaataa-a-cacgaaaa-gatatag-gcacc-cct  
 015-TT GGTTcAGGCTGGATTGCGTCGCAAGCCcACCATAGGCCCTCGCAACG T TCTG-015  
 023-TTA TT GCAAAGTGGATATCGAGGACGGACGT-023  
 6 7209 CTatTCGTACAGGTgTtAAGCGCCcTGTGtTCCAAagCCTCTgCtGCCCCtAAACGtAAgCGGcCC  
 11 7194 CTgCTCGTACAGGTaTaAAGCGCCcAGCTGTgTctAAAgCCTCTaCAGCCCCcAAACGAAaCGTaCC  
 33 7059 CaTCgTCTgCAaaacgcAAAAaggttaaaaAATAAcAcTttGtgtaAttgtgtTAtgtTGTtgtTttg  
 16 7108 CcTCATCTACCTcCTACAAcTgCtAAACG cAAAAACgTAAGctGTAA GtaTTGTATGTA  
 31 7021 CagCATCTACCACTACAcCaGCAAAACGtaAAAAAAC TAAAAaGTAAtgGatgTGTATGTAAaCaT  
 18 7075 CtcCATCTgCCACTAC gtcttC TAAA ccTGccAagCgT  
 con Ct-catcTaC-actacaaa---c--a--t-aat-aa-gtaa-ctg-a-cc-ct-a-c-tgta--tcc-  
 015-CTCCATCTGCCACTAC GTCTTC TAAA CCTGCCAAGCGT-015  
 023-CTGCTCGTACAGGTATAAAGCGCCcAGCTGTGtCTAAGCCCTCTACAGCCCCcAAACGAAaCGTACC-023

6 7277 AAAACtAAAAgGTAATATATGTGT aTaTGTacTGTT  
 11 7262 AAAACcAAAAaGTAATATATGTGTgtcagTGTCTTGTT  
 33 7127 TtcTGtcTAtGTactTtgtgTTGT TGTGTTGTGTTgtTGT  
 16 7167 TgtTGaaTtaGTGT TgtTTGT TGTGT ATATGT  
 31 7088 GTGTctgTatGTGTAtGTGCTTGTgctgtatTGT ATATGTGTGTGTTtgtgtgtTATATA tg  
 18 7113 GTG cGTGTAcGTGC caGgaagtaATATGTGTGTGT gtatataTATATAc  
 con -t-tctataagtgtat-tgtttgtg----tgtGtagtgt-tatgtgtgtgt-----tatata---  
 015-GTG CGTGTACGTGC CAGGAAGTAATATGTGTGTGT GTATATATATATACAT-015  
 023-AAAACCAAAAAGTAATATATGTGTGTGTCAGTGTGTTGTGTT-(023)  
 6 7313 AT  
 11 7302 ATtTaTATG T TGTTGTA gTGTGT  
 33 7167 TTGT TtTtTgTGTATG TGttacaaTgtATgTTATGTTGTATGTTacTGTGTTG  
 16 7199 TTGTATGTgctTGTATG TG CTGTAAaATATTaAGTTGTATGT GTGTTTG  
 31 7150 gTATaTGTATGTTTaTGTATG CgTGTGT aCTGTATATAT GtaTaGTATGT TATGTgTG  
 18 7164 cTATtgtTgTGTTT GTATGcCtGTGTTtgtgtTTGT TgTAT G aTtgcaattgTATG G  
 con -tatttgtatgtttttgtatg-c-tgtgt-tgt-cttgtatatattatgttgtatgtt-gtgtgtttg  
 015-CTATTGTTGTGTTT GTATGTCCTGTGTTTGTGTTTGT TGTAT G ATTGCATTGTATG G-015  
 023- ATTTATATG T TGTTGTA GTGTGT(-023)  
 6 7315 aTATATGT  
 11 7325 aTATGT TtCTTGT atTGTG TATATGT  
 33 7221 T ttTATGTgTaCTTGTtTGTGTGcATGTTcTATGTacttgt  
 16 7248 TATGTATG gtaTAATAAA cacGTGTGTATGT  
 31 7209 TATGTATGctatgtaTGTAAATAAatgtgtataacctgtgtgtgtTGTGTATGTTGcctTataTAc  
 18 7221 TATGTATG gtTGTGTT gtTGTATGTTGTatgTtacTAt  
 con tatgtatg-----tgtaataaaa---ttatgt-ttcttgtt-gtgtgtatggt-tatgta--tat  
 015-TATGTATG GTTGTGTT GTTGTATGTTGTATGTTACTAT-015  
 023- ATATGT TTCTTGT ATTGTG TATATGT(-023)



6 7323 GT gTATGTAcTGT  
 11 7351 GTataTGTtTGTGTATATGT GTAT gtTATGTa TGT  
 33 7262 cagtTtccTGTtTGTGTATATGTtaataaaacattgTgTGTATtTgtTaaActATtTgTATGTa TGT  
 16 7279 GTtTtTaAaTgcTTgtgTAACTATTGT gTcATgcAacATaAaTaaacttatt  
 31 7277 AccctaTtagtaacatacTaTtAcTAtTTtataAACTATTGTtccTActTgtTcctAcTgttCCTgc  
 18 7257 AtttgtTggtatgtggcaTtaaAtaAaaTatgttttgggtTctgTgtgTtaTgtggtTgcgcCCTag  
 con a---tat-tgtttgtgtatat-ataatataaagaactatgtttttatgtaatat--Tatgtactgt  
 015-ATTTGTTGGTATGTGGCATTAAATAAAATATGTTTGTGGTTCGTGTGTTATGTGGTTCGCCCTAG-015  
 023- GTATATGTTTGTGTATATGT GTAT GTTATGTa TGT-023  
 6 7336 TATGT aTATGT GTgTGTGTGTtctGTGTGtaAgtTATTTGTGtAAATGTGTATGTgTGT  
 11 7386 TATGTgtTATGTatGTtGTGTGTtTaaGTGTGT gtATaTATTTGTGGAATGTGTATGTATGTT  
 33 7329 TATGT AtatgggtgtaccTataTgaGTAagGagTTgTATTgcTtGccctacCcTGCATTgc  
 16 7331 gtTTCaacAcctACTaaTgtTgtTgTggtTaTtcAtTGTATaTaAactaTatTtGctACATcCtgTtc  
 31 7345 TccTCCcaAtagtCATgTAcTtAtTctgccTatAaTTAggTgTcacgccaTaGtaAaAgTgtgaca  
 18 7325 TgagtaacAactgtATtTgtgTtTgTggtatgggtgTgcttTgTgggtatataTtgtccTgtattt  
 con tatgtaa-aa-gt-attttgt-tttT-tgtgtgtaaatgtattttatttgt-taa-ttgtatgt-tttt  
 015-TGAGTAACAACCTGTATTTGTGTTTGTGGTATGGGTGTTGCTTGTGGGCTATATATTGTCCTGTATTT-015  
 023-TATGTTGTTATGTATGTTTGTGTGTTTAGTGTGT GTATATATTTGTGGAATGTGTATGTATGTT-023  
 6 7400 TaTGTGCAATAAACAATTAcctcTtgtTcacacCCTGT gACTcAGTGgctgttgcaacgcGTTtTGgT  
 11 7450 TtTGTGCAATAAACAATTA TTatgTgtgtCCTGTTACACcCAGTG actaaGTTgTgT  
 33 7390 aaTGTaCcTAccTtTtccTtATtTccTtATtTgtAGtaCCTACATGTttaGTattgCtttacCtTTTGaca  
 16 7399 ttgtTtTaTATaTactaTAtTtTgTAgcgcCAGgcCCatTTTGTaGCTTCaAcGgaAttCggTGTGcat  
 31 7413 CccGgTccgtTtTTtgcaACTaaAgctacTCCATTTTgaTTTtatGCagCCATTTTAAaTcccTAACC  
 18 7393 CaaGtTataaaacTgcacACcttAcagcaTCCATTTtatccTacaatcctCcaTTTtgcTgtgcAACC  
 con tatgttcaa-aatt-attaccttata-t-tcc-tt-t-acat-cagtg-c-atttta--cgttt-act  
 015-CAAGTTATAAACTGCACACCTTACAGCATCCATTTTATCCTACAATCCTCCATTTTGTGTGCAACC-015  
 023-TTTGTGCAATAAACAATTA TTATGTGTGCTCCTGTACACCCAGTG ACTAAGTTGTGTT-023  
 024-GAATTCGGTTGCAT

6 7466 TTGCACGCGCCCTacacacataagTaATATacaTgcAcaATATATATaTtTttgtTtaaaATACTAT  
 11 7508 TTGCACGCGCCGgTtTgtgttgccTTCATAT TatAtTATATATATtTgTaataTacCTATACTATg  
 5 33 7458 TacTAGTGtCCaTATtgtacaaTTCCtccattTTgTATGcCTAaccgTtTtcggTtACTTgGCAtac  
 16 7467 GcTTTTtTgGCaCAAAaTgTgttTTTTaAaTAgTTCATGtCagcaacTaTgGtTtAaAcTTGTACGT  
 31 7481 GtTTTCGGTTGCAttgTtTaaacaTgctAgTAcacCTATGctgatgcagtaGTTcTGcggTtTtTgGT  
 10 18 7461 GaTTTCGGTTGC ctttggcTTaTGctgTggTttT  
 con -ttt-cgg-ccttat-t-ta-a-ttc-tataa-t-ctatgt-tatat-ttt-tt-T-actttgct-tt  
 015-GATTTTCGGTTGC CTTTGGCTTATGCTGTGGTTTT-015  
 023-TTGCACGCGCCGTTTGTGTGCTTCATAT TATATTATATATATTGTAATATACCTATACATG-023  
 024-GCTTTTGGCACAAATGTGTTTTTTAAATAGTCTATGTGAGCAACTATGTTTTAACTGTGACGT-024  
 15 6 7533 aCttttatAtTTGCAACCGTTTTCGGTTGCCCTTAgCATACACTTtCCaCcAATTGTtAcAAC  
 11 7573 tTACCCccccccCAcTTGCAACCGTTTTCGGTTGCCCTTA CATACACTTAcCTCaAATTGTtAcAAC  
 20 33 7526 aTACCCtaTgaCAtTGGCAGaAcAgTTaaTccTTtCTtCCTGCACTGtgtTgtTcTgTACTtgcgtg  
 16 7535 TTCTTG cTtgCcaTGcgtGccAaAaTcccTgtTTTcCTGaCCTGCACTG cTTgccaACcaTccc  
 31 7549 TTCTTG aaTAcTagTTTTtGcCaacaTCTTggcTgtTagt  
 25 18 7496 cTgCacaatacagtagcgtggcactattgcaaacctTaaTctTTTggGCactgcTcTacaTatTttg  
 con tt-c-ct-tt-catt-gcagcc--tttcg-tt-ctcttatc-T-cact--c-tctttt-tattata-c  
 015-CTGCACAAATACAGTACGCTGGCACTATTGCAAACTTTAATCTTTTGGGCACTGCTCCTACATATTTG-015  
 023-TTACCCCCCCCCACTTGCACCGTTTTCGGTTGCCCTTA CATACACTTACCTCAAAATTTGTATAAC-023  
 024-TTCTTG CTGCCATGCGTGCCAAATCCCTGTTTCTGACCTGCACTG CTTGCCAACCATTC-024  
 30 6 7597 GTGTTTccTctTAATCCtATATattTTGTG CcAGGTACAcATTGCCCTGCCAAGTtgCTTGCCAA  
 11 7640 GTGTTTtgTACTAATCCcATAT gTTGTgtgcCAAGGTACAtATTGCCCTGCCAAGTatCTTGCCAA  
 33 7594 caTTggcaTACatAcCCtAtgacatTgGCagaaCagTtAAcctTTtCTTtCTgcacTgtgTtTgtc  
 35 16 7598 aTTgTtttTtACACTgCacTatgtgcaActActgAaTCACTaTgTaCATTgtgTCataTAAaTaaaT  
 31 7589 tTCCtGccTaACACacCTTgccaaCATATAAtccAgTCCaacTtTGCAATTataCtATgAAAtCatgtT  
 18 7564 aaCaattggcgCgCctCTTtggcgCATATAA ggCgcaccTGgtATTA gtcATtttcCtgtcc  
 40 con -t-ttt--ta-ca-tcCtatat---tt-taa-ccaa-g-aca--Ttgc-tt-c--aatt--ttt--a-  
 015-AACAATTGGCGCGCCTCTTTGGCGCATATAA GGCGCACCTGGTATTA GTCATTTTCTGTGCC-015  
 023-GTGTTTTGTACTAATCCCATAT G-023  
 024-ATTGTTTTTACACTGCACTATGTGCAACTACTGAATCACTATGTACATTGTGTCATATAAAATAAAT-024

6 7662 gtgcatcatatcctgccaaCCACACACCTGGCgcCAGGGtGCGGTATTGC cTtactcATAA  
 11 7706 CAACACACCTGGC CAGGGcGCGGTATTGCATgAcTaaTgTacaATAA  
 5 33 7662 tgtacTtgctgcAttgacTCatataCatGCAGtgcaATTgcaaAaTactTAATTgtacTAatAgtT  
 16 7666 cAcTaTgcgcCAACGcctTacatACcgCtgTgAGgcacATatTTTggcTTgTtTAactAACcTAAT  
 31 7657 tGtTtsaaTACAAcTgtagttcaACtATgtgTcatgcAcaTATATTataTTaTCCTACAcAcCTTAAA  
 10 18 7626 aGgTgcgcTACAAC aATTgcTtgcatAacTATAT ccactcCCTA AgtaaTAAAA  
 con tg-tatg-tacaacgccatc-a-acaactgg-agca-aatt-tata-t-cttt-cta-a--actaaaa  
 BE31-XXAGGCACAXAXXXX-BE31 hpv16+18+33  
 015-AGGTGCGCTACAAc AATTGCTTGcATAACTATAT CCACTCCCTA AGTAATAAAA-015  
 024-CACTATGCGCCAACGCCCTTACATACCGCTGTAGGCACATATTTTGGCTTGTTTAACTAACCTAAT-024  
 15 6 7723 ACCTGTC TTTGTgttAtAcTtTTaTGcAcTGtAGCCAAcTcTTAAAAGCATTtTTGGCTTgTAGCa  
 11 7753 ACCTGTCGgTTTGT ACAaTgTTgTgATTGCAGCCAAAggTTAAAAGCATTtTTGGCTTgTAGCt  
 20 33 7730 TaCacATGcTTTtaggcACATAtTTTtactTTaCtttCAAaccTTAAgtGCAGTTTGGCTT aCa  
 16 7734 TgCATATTtGGCAtAaggTTTAAacTTcTAaggCcAaCtAAatgTcAcacctAGTTCaTaCaTgaActg  
 31 7725 CTGCTTTTAGGCACATATTTT GTagaTTATctaTatCctTgATTGCAgtgcTGGCTTtgcacAtgt  
 18 7680 CTGCTTTTAGGCACATATTTTaGttgtTtTtacTtaagcTaATTGCAtactTGGCTT  
 25 con --c-tttta---atataat-tagtttt-tattgct--caaa--tTaaa-gcattt-t-gcttgtagc-  
 BE31-XXAGGCACAXAXXXX-BE31 hpv16+18+33  
 BE31-XXAGGCACAXAXXXX-BE31 hpv16+18+33  
 015-CTGCTTTTAGGCACATATTTTAGTTTGTtTTTACTTAAGCTAATTGCATACTTGGCTT-(015)  
 024-TGCATATTGGCATAAGGTTTAACTTCTAAGGCCAACTAAATGTCACCCTAGTTCATACATGAACCTG-024  
 30 6 7789 GcACATTTTTtTgCtCTTAcTgTtTggTatACAATAaCataAAAAATGAGTAACCTAAGGTCACACACC  
 11 7818 GAACATTTTGTACcCTTAgTaTaTtaTgcACAATAcCcacAAAAATGAGTAACCTAAGGTCACACACC  
 33 7795 cAAAttgcTTGTATgCcaAactaTgcccTTGTAAAAgtgagtcActacctgttTaTtAccaGGTGTGga  
 35 16 7802 TgtAAagGTTAgTcaTAcATgTTCATTGTAAAA cTgcAcatgGGTGTGtg  
 31 7792 TtaAAActGccAaggTTgtgTcaTgCATTaTaAATAagttgTatgttactcaTATAATtaATgCatAt  
 18 7738 gtacaactacTTtcaTgtcccaAcacTctgTctacccTtaacatgaacTATAAT ATGaCtaAg  
 40 con -aa-atTTTT-tact-ttatt-tt-a-tttaaaaaac-gtaaa-tg--tat--taagga-g--ta--  
 015- GTACAACTACTTTcATGTCCAAcATtCTGTCTACCCCTTAACATGAACCTATAAT ATGACTAAG-015  
 024-TGTAAAGGTTAGTCATACATTGTTcATTGTAAAA CTGCACATGGGTGTGTG-024  
 45  
 50  
 55

## 25

45  
50

2. A composition according to claim 1 for amplifying the DNA of human papilloma virus type 16 present in a test sample, said composition comprising a set of four oligonucleotide probes, said probe sets being selected from the group consisting of the following oligonucleotide sets:  
LCR5 (SEQ ID Nos. 81,82,83 and 84) and LCR8 (SEQ ID Nos. 93, 94, 95 and 96).

3. A composition according to claim 1 for amplifying the DNA of human papilloma virus type 18 present in a test sample, said composition comprising a set of four oligonucleotide probes, said probe sets being selected from the group consisting of the following oligonucleotide sets:  
LCR 6 (SEQ ID Nos. 85,86,87 and 88) and LCR 7 (SEQ ID Nos. 89,90,91 and 92).

4. A kit for detecting the presence of human papilloma virus DNA in a test sample, comprising:  
a composition according to any of claims 1 to 3; and  
further comprising a ligase.

5. A kit according to claim 4, wherein said ligase is thermostable.

6. A composition useful in PCR for amplifying the DNA of human papilloma virus present in a test sample, said composition comprising:

a first nucleic acid primer of sense direction, capable of hybridizing to the antisense strand of HPV DNA, said primer having from 10 to about 30 nucleotides in length and having a sequence selected from the group consisting of the following sequences:

SEQ ID No.

1 CAGATGTCTC TGTGGCGGCC TAGTG.

6 GAATTAGTTA GACCATTTAA AAG.

7 GGGGAAACAC CAGAATGGAT A.

81 GCTGCAAACA ACTATACATG ATATAA.

85 CTTCACTGCA AGACATAGAA ATAA.

89 TATATTGCAA GACAGTATTG GAAC and

93 GTATGGAACA ACATTAGAAC AGCA; and

a second nucleic acid primer of antisense direction, capable of hybridizing to the sense strand of HPV DNA, said primer having from 10 to about 30 nucleotides in length and having a sequence selected from the group consisting of the following sequences:

SEQ ID No.

5 AGGTGTCAGG AAAACCAAAT TTATT.

84 TGCTTGCACT ACACACATTG TAATA.

88 TACTGTCTTG CAATATACAC AGG.

92 AATGCAAATT CAAATACCTC TGTA and

96 AAATCACACA ACGGTTTGTT GTAT;

provided said first and second primers hybridize to their respective antisense and sense strands at locations such that their 3' ends do not overlap and, in the direction of extension, the 5' ends of said primers are spaced further apart than the 3' ends of said primers.

7. A composition according to claim 6 wherein said first and second primers are selected from the following pairs of oligonucleotide sequences ( identified by Sequence ID No.):  
1 and 5, 6 and 5, 7 and 5, 81 and 84,  
85 and 88, 89 and 92, and 93 and 96.

8. A kit for detecting the presence of human papilloma virus DNA in a test sample, comprising:  
a composition according to claim 6 or 7; and  
further comprising a polymerase.

9. A kit according to claim 8 wherein said polymerase is thermostable.

10. A consensus oligonucleotide for hybridizing human papilloma virus types 6, 11, 16, 18, 31, 33 and 61, which oligonucleotide comprises from about 10 to about 60 nucleotides in length and is selected from the group of sequences consisting of:

SEQ ID No.

1	CAGATGTCTC	TGTGGCGGCC	TAGTG,
5	AGGTGTCAGG	AAAACCAAAT	TTATT,
6	GAATTAGTTA	GACCATTTAA	AAG and
7	GGGGAACAC	CAGAATGGAT	A;

and their complements.

11. A type-specific oligonucleotide for determining the presence of human papilloma virus type 16, having a sequence selected from the group consisting of:

SEQ ID No.

81	GCTGCAAACA	ACTATACATG	ATATAA,
82	TTATATCATG	TATAGTTGTT	TGCAGC,
83	TATTAGAATG	TGTGTACTGC	AAGCA,
84	TGCTTGCACT	ACACACATTC	TAATA,
93	GTATGGAACA	ACATTAGAAC	AGCA,
94	TGCTGTTCTA	ATGTTGTTCC	ATAC,
95	ATACAACAAA	CCGTTGTGTG	ATTT and
96	AAATCACACA	ACGGTTTGTG	GTAT;

and their complements.

12. A type-specific oligonucleotide for determining the presence of human papilloma virus type 18, having a sequence selected from the group consisting of: SEQ ID No.

SEQ ID No.

85	CTTCACTGCA	AGACATAGAA	ATAA,
86	TTATTTCTAT	GTCCTGCACT	GAA,
87	CCTGTGTATA	TTGCAAGACA	GTAT,
88	TACTGTCTTG	CAATATACAC	AGG,
89	TATATTGCAA	GACAGTATTG	GAAC,
90	GTTCCAATAC	TGCTTTGCAA	TTTA,
91	TTACAGAGGT	ATTTGAATTT	GCATT and
92	AATGCAAATT	CAATACCTC	TGTAA;

and their complements.

13. A method for determining the presence of any human papilloma virus in a test sample, comprising:

- a. hybridizing DNA in the test sample with at least one consensus oligonucleotide selected from the group of claim 10, said oligonucleotide being conjugated to a signal generating compound capable of producing a detectable signal; and
- b. determining the presence of human papilloma virus by detecting the signal generated.

14. A method for determining the presence of human papilloma virus type 16 in a test sample, comprising:

- a. hybridizing DNA in the test sample with at least one oligonucleotide selected from the group of claim 11, said oligonucleotide being conjugated to a signal generating compound capable of producing a detectable signal; and
- b. determining the presence of human papilloma virus by detecting the signal generated.

15. A method for determining the presence of human papilloma virus type 18 in a test sample, comprising:

- a. hybridizing DNA in the test sample with at least one oligonucleotide selected from the group of claim 12, said oligonucleotide being conjugated to a signal generating compound capable of producing a detectable signal; and
- b. determining the presence of human papilloma virus by detecting the signal generated

16. A method according to any of claims 13-15, further comprising a step of amplification prior to or concurrent with said hybridizing step.

17. A method according to claim 16, wherein said amplification step comprises PCR or LCR.

**Claims for the following Contracting States : ES**

1. A composition useful in LCR for amplifying the DNA of human papilloma virus present in a test sample, said composition comprising a set of four oligonucleotide probes, said probe sets being selected from the group consisting of the following oligonucleotide sets:

LCR5: SEQ ID No.

81	GCTGCAAACA	ACTATACATG	ATATAA,
82	TTATATCATG	TATAGTTGTT	TGCAGC,
83	TATTAGAATG	TGTGTACTGC	AAGCA,
84	TGCTTGCAGT	ACACACATTG	TAATA;

LCR6: SEQ ID No.

85	CTTCACTGCA	AGACATAGAA	ATAA,
86	TTATTTCTAT	GTCTTGCAGT	GAA,
87	CCTGTGTATA	TTGCAAGACA	GTAT,
88	TACTGTCTTG	CAATATACAC	AGG;

LCR7: SEQ ID No.

89	TATATTGCAA	GACAGTATTG	GAAC,
90	GTTCCAATAC	TGTCTTGCAA	TTTA,
91	TTACAGAGGT	ATTTGAATTT	GCATT,
92	AATGCAAATT	CAAATACCTC	TGTAA; and

LCR8: SEQ ID No.

93	GTATGGAACA	ACATTAGAAC	AGCA,
94	TGCTGTTCTA	ATGTTGTTCC	ATAC,
95	ATACAACAAA	CCGTTGTGTG	ATTT,
96	AAATCACACA	ACGGTTTGTG	GTAT.

2. A composition according to claim 1 for amplifying the DNA of human papilloma virus type 16 present in a test sample, said composition comprising a set of four oligonucleotide probes, said probe sets being selected from the group consisting of the following oligonucleotide sets:

LCR5 (SEQ ID Nos. 81,82,83 and 84) and LCR8 (SEQ ID Nos. 93, 94, 95 and 96).

3. A composition according to claim 1 for amplifying the DNA of human papilloma virus type 18 present in a test sample, said composition comprising a set of four oligonucleotide probes, said probe sets being selected from the group consisting of the following oligonucleotide sets:

LCR6 (SEQ ID Nos. 85,86,87 and 88) and LCR 7 (SEQ ID Nos. 89,90,91 and 92).

4. A kit for detecting the presence of human papilloma virus DNA in a test sample, comprising:

EP 0 477 972 B1

a composition according to any of claims 1 to 3; and further comprising a ligase.

5. A kit according to claim 4, wherein said ligase is thermostable.

6. A composition useful in PCR for amplifying the DNA of human papilloma virus present in a test sample, said composition comprising:

a first nucleic acid primer of sense direction, capable of hybridizing to the antisense strand of HPV DNA, said primer having from 10 to about 30 nucleotides in length and having a sequence selected from the group consisting of the following sequences:

SEQ ID No.

1 CAGATGTCTC TGTGGCGGCC TAGTG.

6 GAATTAGTTA GACCATTAA AAG.

7 GGGGAAACAC CAGAATGGAT A.

81 GCTGCAAACA ACTATACATG ATATAA.

85 CTTCACTGCA AGACATAGAA ATAA.

89 TATATTGCAA GACAGTATTG GAAC and

93 GTATGGAACA ACATTAGAAC AGCA; and

a second nucleic acid primer of antisense direction, capable of hybridizing to the sense strand of HPV DNA, said primer having from 10 to about 30 nucleotides in length and having a sequence selected from the group consisting of the following sequences:

SEQ ID No.

5 AGGTGTCAGG AAAACCAAAT TTATT.

84 TGCTTGCACT ACACACATTC TAATA.

88 TACTGTCTTG CAATATACAC AGG.

92 AATGCAAATT CAAATACCTC TGTA and

96 AAATCACACA ACGGTTTGT GTAT;

provided said first and second primers hybridize to their respective antisense and sense strands at locations such that their 3' ends do not overlap and, in the direction of extension, the 5' ends of said primers are spaced further apart than the 3' ends of said primers.

7. A composition according to claim 6 wherein said first and second primers are selected from the following pairs of oligonucleotide sequences (identified by Sequence ID No.):  
1 and 5, 6 and 5, 7 and 5, 81 and 84,  
85 and 88, 89 and 92, and 93 and 96.

8. A kit for detecting the presence of human papilloma virus DNA in a test sample, comprising:  
a composition according to claim 6 or 7; and  
further comprising a polymerase.

9. A kit according to claim 8 wherein said polymerase is thermostable.

10. A method for determining the presence of any human papilloma virus in a test sample, comprising:

a. hybridizing DNA in the test sample with at least one consensus oligonucleotide selected from the group of sequences consisting of:



SEQ ID No.

1 CAGATGTCTC TGTGGCGGCC TAGTG,  
 5 AGGTGTCAGG AAAACCAAAT TTATT,  
 6 GAATTAGTTA GACCATTTAA AAG and  
 7 GGGGAAACAC CAGAATGGAT A;

and their complements,

said oligonucleotide being conjugated to a signal generating compound capable of producing a detectable signal; and

b. determining the presence of human papilloma virus by detecting the signal generated.

11. A method for determining the presence of human papilloma virus type 16 in a test sample, comprising:

a. hybridizing DNA in the test sample with at least one oligonucleotide selected from the group of sequences consisting of:

SEQ ID No.

81 GCTGCAAACA ACTATACATG ATATAA,  
 82 TTATATCATG TATAGTTGTT TGCAGC,  
 83 TATTAGAATG TGTGTACTGC AAGCA,  
 84 TGCTTGCACT ACACACATTC TAATA,  
 93 GTATGGAACA ACATTAGAAC AGCA,  
 94 TGCTGTTCTA ATGTTGTTC ATAC,  
 95 ATACAACAAA CCGTTGTGTG ATTT and  
 96 AAATCACACA ACGGTTTGT GTAT;

and their complements, said oligonucleotide being conjugated to a signal generating compound capable of producing a detectable signal; and

b. determining the presence of human papilloma virus by detecting the signal generated.

12. A method for determining the presence of human papilloma virus type 18 in a test sample, comprising:

a. hybridizing DNA in the test sample with at least one oligonucleotide selected from the group of sequences consisting of:

SEQ ID No.

85 CTTCACTGCA AGACATAGAA ATAA,  
 86 TTATTTCTAT GTCITGCAGT GAA,  
 87 CCTGTGTATA TTGCAAGACA GTAT,  
 88 TACTGTCTTG CAATATACAC AGG,  
 89 TATATTGCAA GACAGTATTG GAAC,  
 90 GTTCCAATAC TGTCTTGCAA TTTA,  
 91 TTACAGAGGT ATTTGAATTT GCATT and  
 92 AATGCAAATT CAAATACCTC TGTA;

and their complements,

said oligonucleotide being conjugated to a signal generating compound capable of producing a detectable signal; and

b. determining the presence of human papilloma virus by detecting the signal generated.

13. A method according to any of claims 10-12, further comprising a step of amplification prior to or concurrent with said hybridizing step.

14. A method according to claim 13, wherein said amplification step comprises PCR or LCR.

## Patentansprüche

Patentansprüche für folgende Vertragsstaaten : AT, BE, CH, LI, DE, DK, FR, GB, GR, IT, NL, SE

1. Zusammensetzung, die für die LCR ("ligase chain reaction", Ligasekettenreaktion) zur Vervielfachung der DNA des humanen Papillomavirus nützlich ist, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung einen Satz von vier Oligonukleotidsonden umfaßt, wobei die SONDENSÄTZE aus der Gruppe gewählt sind, die aus den folgenden Oligonukleotidsätzen besteht:

## LCR5: SEQ ID N r

```

81      GCTGCAAACA ACTATACATG ATATAA.
82      TTATATCATG TATAGTTGTT TGCAGC.
83      TATTAGAATG TGTGTACTGC AAGCA.
84      TGCTTGCAGT ACACACATTG TAATA.

```

## LCR6: SEQ ID N r

```

85      CTTCACTGCA AGACATAGAA ATAA.
86      TTATTTCTAT GTCTTGCAGT GAA.
87      CCTGTGTATA TTGCAAGACA GTAT.
88      TACTGTCTTG CAATATACAC AGG.

```

## LCR7: SEQ ID N r

```

89      TATATTGCAA GACAGTATTG GAAC.
90      GTTCCAATAC TGTCTTGCAA TTTA.
91      TTACAGAGGT ATTTGAATTT GCATT.
92      AATGCAAATT CAAATACCTC TGTA. und

```

## LCR8: SEQ ID N r

```

93      GTATGGAACA ACATTAGAAC AGCA.
94      TGCTGTTCTA ATGTTGTTCC ATAC.
95      ATACAACAAA CCGTTGTGTG ATTT.
96      AAATCACACA ACGGTTTGTG GTAT.

```

2. Zusammensetzung nach Anspruch 1 zur Vervielfachung der DNA des humanen Papillomavirus Typ 16, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung einen Satz von vier Oligonukleotidsonden umfaßt, wobei die SONDENSÄTZE aus der Gruppe gewählt sind, die aus den folgenden Oligonukleotidsätzen besteht: LCR5 (SEQ ID Nrn 81, 82, 83 und 84) und LCR8 (SEQ ID Nrn 93, 94, 95 und 96)
3. Zusammensetzung nach Anspruch 1 zur Vervielfachung der DNA des humanen Papillomavirus Typ 18, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung einen Satz von vier Oligonukleotidsonden umfaßt, wobei die SONDENSÄTZE aus der Gruppe gewählt sind, die aus den folgenden Oligonukleotidsätzen besteht: LCR6 (SEQ ID Nrn 85, 86, 87 und 88) und LCR7 (SEQ ID Nrn 89, 90, 91 und 92).
4. Kit zum Nachweis der Anwesenheit der DNA des humanen Papillomavirus in einer Testprobe, das folgendes umfaßt:  
eine Zusammensetzung nach einem der Ansprüche 1 bis 3, und des weiteren eine Ligase.
5. Kit nach Anspruch 4, worin die Ligase thermostabil ist.
6. Zusammensetzung, die bei der PCR ("polymerase chain reaction" Polymerasekettenreaktion) zur Vervielfachung der DNA des humanen Papillomavirus nützlich ist, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung folgendes umfaßt:

# EP 0 477 972 B1

einen ersten Nukleinsäureprimer, der zur Richtung gleichläufig ist, welcher zur Hybridisierung an den gegenläufigen Strang der HPV-DNA befähigt ist, wobei der Primer 10 bis ungefähr 30 Nukleotide lang ist und eine Sequenz aufweist, die aus der Gruppe gewählt ist, die aus den folgenden Sequenzen besteht:

5  
1 CAGATGTCTC TGTGGCGGCC TAGTG,  
6 GAATTAGTTA GACCATTTAA AAG,  
10 7 GGGGAAACAC CAGAATGGAT A,  
81 GCTGCAAACA ACTATACATG ATATAA,  
85 CTTCACTGCA AGACATAGAA ATAA,  
89 TATATTGCAA GACAGTATTG GAAC und  
15 93 GTATGGAACA ACATTAGAAC AGCA; und

einen zweiten Nukleinsäureprimer, der zur Richtung gegenläufig ist, welcher zur Hybridisierung an den gleichläufigen Strang der HPV-DNA befähigt ist, wobei der Primer 10 bis ungefähr 30 Nukleotide lang ist und eine Sequenz aufweist, die aus der Gruppe gewählt ist, die aus den folgenden Sequenzen besteht:

20 SEO ID Nr

25 5 AGGTGTCAGG AAAACCAAAT TTATT,  
84 TGCTTGCAGT ACACACATTTC TAATA,  
88 TACTGTCTTG CAATATACAC AGG,  
92 AATGCAAATT CAAATACCTC TGTA und  
30 96 AAATCACACA ACGGTTTGT GTAT;

vorausgesetzt, daß der erste und der zweite Primer an ihre jeweiligen gleich- und gegenläufigen Stränge an solchen Stellen hybridisieren, daß ihre 3'-Enden nicht überlappen, und daß die 5'-Enden der Primer in Verlängerungsrichtung weiter räumlich abgesetzt sind als die 3'-Enden der Primer.

- 35 7. Zusammensetzung nach Anspruch 6, worin der erste und zweite Primer aus den folgenden Paaren von Oligonukleotidsequenzen (die durch die Sequenz ID Nr bezeichnet sind) gewählt sind:  
1 und 5, 6 und 5, 7 und 5, 81 und 84,  
85 und 88, 89 und 92, und 93 und 96.
- 40 8. Kit zum Nachweis der Anwesenheit der DNA des humanen Papillomavirus in einer Testprobe, das folgendes umfaßt:  
eine Zusammensetzung nach Anspruch 6 oder 7 und des weiteren eine Polymerase.
- 45 9. Kit nach Anspruch 8, worin die Polymerase thermostabil ist.
10. Consensus-Oligonukleotid zur Hybridisierung der humanen papillomaviren Typ 6, 11, 16, 18, 31, 33 und 61, wobei das Oligonukleotid ungefähr 10 bis ungefähr 60 Oligonukleotide lang ist und aus der Gruppe von Sequenzen gewählt ist, die aus folgendem besteht:

50 SEO ID Nr

55 1 CAGATGTCTC TGTGGCGGCC TAGTG,  
5 AGGTGTCAGG AAAACCAAAT TTATT,  
6 GAATTAGTTA GACCATTTAA AAG und  
7 GGGGAAACAC CAGAATGGAT A;

und aus deren Komplementen.

11. Typ-spezifisches Oligonukleotid zur Bestimmung der Anwesenheit des humanen Papillomavirus Typ 16, das eine Sequenz aufweist, die aus der Gruppe gewählt ist, die aus folgendem besteht:

5

SEQ ID Nr

	81	GCTGCAAACA	ACTATACATG	ATATAA,
10	82	TTATATCATG	TATAGTTGTT	TGCAGC,
	83	TATTAGAATG	TGTGTACTGC	AAGCA,
	84	TGCTTGCACT	ACACACATTC	TAATA,
	93	GTATGGAACA	ACATTAGAAC	AGCA,
	94	TGCTGTTCTA	ATGTTGTTCC	ATAC,
15	95	ATACAACAAA	CCGTTGTGTG	ATTT
	96	AAATCACACA	ACGGTTTGTT	GTAT; und

und aus deren Komplementen.

12. Typ-spezifisches Oligonukleotid zur Bestimmung der Anwesenheit des humanen Papillomavirus Typ 18, das eine Sequenz aufweist, die aus der Gruppe gewählt ist, die aus folgendem besteht:

25

SEQ ID Nr

	85	CTTCACTGCA	AGACATAGAA	ATAA,
	86	TTATTTCTAT	GTCTTGCACT	GAA,
	87	CCTGTGTATA	TTGCAAGACA	GTAT,
30	88	TACTGTCTTG	CAATATACAC	AGG,
	89	TATATTGCAA	GACAGTATTG	GAAC,
	90	GTTCCAATAC	TGTCTTGCAA	TTTA,
	91	TTACAGAGGT	ATTTGAATTT	GCATT und
	92	AATGCAAATT	CAAATACCTC	TGTAA;

35 und aus deren Komplementen.

13. Verfahren zur Bestimmung der Anwesenheit irgendeines humanen Papillomavirus in einer Testprobe, das folgendes umfaßt:

- 40 a. Hybridisieren der DNA in der Testprobe mit wenigstens einem Consensus-Oligonukleotid, das aus der Gruppe nach Anspruch 10 gewählt ist, wobei das Oligonukleotid an eine signalerzeugende Verbindung konjugiert ist, die zur Erzeugung eines nachweisbaren Signals befähigt ist, und  
b. Bestimmen der Anwesenheit des humanen Papillomavirus, indem das erzeugte Signal nachgewiesen wird.

- 45 14. Verfahren zur Bestimmung der Anwesenheit des humanen Papillomavirus Typ 16 in einer Probe, das folgendes umfaßt:

- 50 a. Hybridisieren der DNA in der Testprobe mit wenigstens einem Oligonukleotid, das aus der Gruppe nach Anspruch 11 gewählt ist, wobei das Oligonukleotid an eine signalerzeugende Verbindung konjugiert ist, die zur Erzeugung eines nachweisbaren Signals befähigt ist, und  
b. Bestimmen der Anwesenheit des humanen Papillomavirus, indem das erzeugte Signal nachgewiesen wird.

- 55 15. Verfahren zur Bestimmung der Anwesenheit des humanen Papillomavirus Typ 18 in einer Testprobe, das folgendes umfaßt:

- a. Hybridisieren der DNA in der Testprobe mit wenigstens einem Oligonukleotid, das aus der Gruppe nach Anspruch 12 gewählt ist, wobei das Oligonukleotid an eine signalerzeugende Verbindung konjugiert ist, die zur Erzeugung eines nachweisbaren Signals befähigt ist, und

b. Bestimmen der Anwesenheit des humanen Papillomavirus, indem das erzeugte Signal nachgewiesen wird.

16. Verfahren nach einem der Ansprüche 13-15, das des weiteren einen Vervielfachungsschritt umfaßt, der vor oder in Konkurrenz mit dem Hybridisierungsschritt stattfindet.

17. Verfahren nach Anspruch 16, worin der Vervielfachungsschritt PCR oder LCR umfaßt.

**Patentansprüche für folgenden Vertragsstaat : ES**

1. Zusammensetzung, die für die LCR ("ligase chain reaction", Ligasekettenreaktion) zur Vervielfachung der DNA des humanen Papillomavirus nützlich ist, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung einen Satz von vier Oligonukleotidsonden umfaßt, wobei die SONDENSÄTZE aus der Gruppe gewählt sind, die aus den folgenden Oligonukleotidsätzen besteht:

LCR5: SEQ ID N<sup>o</sup>

81	GCTGCAAACA	ACTATACATG	ATATAA,
82	TTATATCATG	TATAGTTGTT	TGCAGC,
83	TATTAGAATG	TGTGTACTGC	AAGCA,
84	TGCTTGCACT	ACACACATTC	TAATA;

LCR6: SEQ ID N<sup>o</sup>

85	CTTCACTGCA	AGACATAGAA	ATAA,
86	TTATTTCTAT	GTCTTGCACT	GAA,
87	CCTGTGTATA	TTGCAAGACA	GTAT,
88	TACTGTCTTG	CAATATACAC	AGG;

LCR7: SEQ ID N<sup>o</sup>

89	TATATTGCAA	GACAGTATTG	GAAC,
90	GTTCCAATAC	TGTCTTGCAA	TTTA,
91	TTACAGAGGT	ATTTGAATTT	GCATT,
92	AATGCAAATT	CAAATACCTC	TGTAA;

LCR8: SEQ ID N<sup>o</sup>

93	GTATGGAACA	ACATTAGAAC	AGCA,	und
94	TGCTGTTCTA	ATGTTGTTCC	ATAC,	
95	ATACAACAAA	CCGTTGTGTG	ATTT,	
96	AAATCACACA	ACGGTTTGTT	GTAT,	

2. Zusammensetzung nach Anspruch 1 zur Vervielfachung der DNA des humanen Papillomavirus Typ 16, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung einen Satz von vier Oligonukleotidsonden umfaßt, wobei die SONDENSÄTZE aus der Gruppe gewählt sind, die aus den folgenden Oligonukleotidsätzen besteht: LCR5 (SEQ ID N<sup>o</sup> 81, 82, 83 und 84) und LCR8 (SEQ ID N<sup>o</sup> 93, 94, 95 und 96).

3. Zusammensetzung nach Anspruch 1 zur Vervielfachung der DNA des humanen Papillomavirus TYP 18, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung einen Satz von vier Oligonukleotidsonden umfaßt, wobei die SONDENSÄTZE aus der Gruppe gewählt sind, die aus den folgenden Oligonukleotidsätzen besteht: LCR6 (SEQ ID N<sup>o</sup> 85, 86, 87 und 88) und LCR7 (SEQ ID N<sup>o</sup> 89, 90, 91 und 92).

4. Kit zum Nachweis der Anwesenheit der DNA des humanen Papillomavirus in einer Testprobe, das folgendes umfaßt:  
eine Zusammensetzung nach einem der Ansprüche 1 bis 3, und des weiteren eine Ligase.

5. Kit nach Anspruch 4, worin die Ligase thermostabil ist.

6. Zusammensetzung, die bei der PCR ("polymerase chain reaction" polymerasekettenreaktion) zur Vervielfachung der DNA des humanen Papillomavirus nützlich ist, der in einer Testprobe vorhanden ist, wobei die Zusammensetzung folgendes umfaßt:

5 einen ersten Nukleinsäureprimer, der zur Richtung gleichläufig ist, welcher zur Hybridisierung an den gegenläufigen Strang der HPV-DNA befähigt ist, wobei der Primer 10 bis ungefähr 30 Nukleotide lang ist und eine Sequenz aufweist, die aus der Gruppe gewählt ist, die aus den folgenden Sequenzen besteht:

SEQ ID Nr

10

1 CAGATGTCTC TGTGGCGGCC TAGTG.

15

6 GAATTAGTTA GACCATTAA AAG,  
7 GGGGAAACAC CAGAATGGAT A,

20

81 GCTGCAAACA ACTATACATG ATATAA,  
85 CTTCACAGCA AGACATAGAA ATAA,  
89 TATATTGCAA GACAGTATTG GAAC und  
93 GTATGGAACA ACATTAGAAC AGCA; und

25

einen zweiten Nukleinsäureprimer, der zur Richtung gegenläufig ist, welcher zur Hybridisierung an den gleichläufigen Strang der HPV-DNA befähigt ist, wobei der Primer 10 bis ungefähr 30 Nukleotide lang ist und eine Sequenz aufweist, die aus der Gruppe gewählt ist, die aus den folgenden Sequenzen besteht:

SEQ ID Nr

30

5 AGGTGTCAGG AAAACCAAAT TTATT.

35

84 TGCTTGCACT ACACACATTC TAATA,  
88 TACTGTCTTG CAATATACAC AGG,  
92 AATGCAAAT CAAATACCTC TGTA und  
96 AAATCACACA ACGGTTTGT GTAT;

40

vorausgesetzt, daß der erste und der zweite Primer an ihre jeweiligen gleich- und gegenläufigen Stränge an solchen Stellen hybridisieren, daß ihre 3'-Enden nicht überlappen, und daß die 5'-Enden der Primer in Verlängerungsrichtung weiter räumlich abgesetzt sind als die 3'-Enden der Primer.

45

7. Zusammensetzung nach Anspruch 6, worin der erste und zweite Primer aus den folgenden Paaren von Oligonukleotidsequenzen (die durch die Sequenz ID Nr bezeichnet sind) gewählt sind:

1 und 5, 6 und 5, 7 und 5, 81 und 84,  
85 und 88, 89 und 92, und 93 und 96.

50

8. Kit zum Nachweis der Anwesenheit der DNA des humanen Papillomavirus in einer Testprobe, das folgendes umfaßt:

eine Zusammensetzung nach Anspruch 6 oder 7, und des weiteren eine Polymerase.

9. Kit nach Anspruch 8, worin die Polymerase thermostabil ist.

55

10. Verfahren zur Bestimmung der Anwesenheit irgendeines humanen papillomavirus in einer Testprobe, das folgendes umfaßt:

a. Hybridisieren der DNA in der Testprobe mit wenigstens einem Consensus-Oligonukleotid, das aus der Gruppe von Sequenzen gewählt ist, die aus folgendem besteht:

SEQ ID Nr

5

1 CAGATGTCTC TGTGGCGGCC TAGTG,  
5 AGGTGTCAGG AAAACCAAAT TTATT  
10 6 GAATTAGTTA GACCATTAA AAG und  
7 GGGGAAACAC CAGAATGGAT A;

und aus deren Komplementen,  
wobei das Oligonukleotid an eine signalerzeugende Verbindung konjugiert ist, die zur Erzeugung eines nach-  
weisbaren Signals befähigt ist, und  
15 b. Bestimmen der Anwesenheit des humanen Papillomavirus, indem das erzeugte Signal nachgewiesen wird.

11. Verfahren zur Bestimmung der Anwesenheit des humanen Papillomavirus Typ 16 in einer Probe, das folgendes  
umfaßt:

20

a. Hybridisieren der DNA in der Testprobe mit wenigstens einem Oligonukleotid, das aus der Gruppe von  
Sequenzen gewählt ist, die aus folgendem besteht:

SEQ ID Nr

25

30

81 GCTGCAAACA ACTATACATG ATATAA,  
82 TTATATCATG TATAGTTGTT TGCAGC,  
35 83 TATTAGAATG TGTGTACTGC AAGCA,  
84 TGCTTGCACT ACACACATTC TAATA,  
93 GTATGGAACA ACATTAGAAC AGCA,  
94 TGCTGTTCTA ATGTTGTTCC ATAC,  
95 ATACAACAAA CCGTTGTGTG ATTT und  
40 96 AAATCACACA ACGGTTTGTT GTAT;

und aus deren Komplementen,  
wobei das Oligonukleotid an eine signalerzeugende Verbindung konjugiert ist, die zur Erzeugung eines nach-  
weisbaren Signals befähigt ist, und  
45 b. Bestimmen der Anwesenheit des humanen Papillomavirus, indem das erzeugte Signal nachgewiesen wird.

12. Verfahren zur Bestimmung der Anwesenheit des humanen Papillomavirus Typ 18 in einer Testprobe, das folgendes  
umfaßt:

50

a. Hybridisieren der DNA in der Testprobe mit wenigstens einem Oligonukleotid, das aus der Gruppe von  
Sequenzen gewählt ist, die aus folgendem besteht:

55

SEQ ID Nr

5  
85 CTTCACTGCA AGACATAGAA ATAA,  
86 TTATTTCTAT GTCTTGCAGT GAA,  
87 CCTGTGTATA TTGCAAGACA GTAT,  
88 TACTGTCTTG CAATATACAC AGG,  
89 TATATTGCAA GACAGTATTG GAAC,  
90 GTTCCAATAC TGTCTTGCAA TTTA,  
10 91 TTACAGAGGT ATTTGAATTT GCATT und  
92 AATGCAAATT CAAATACCTC TGTA;

und aus deren Komplementen,  
15 wobei das Oligonukleotid an eine signalerzeugende Verbindung konjugiert ist, die zur Erzeugung eines nachweisbaren Signals befähigt ist, und  
b. Bestimmen der Anwesenheit des humanen Papillomavirus, indem das erzeugte Signal nachgewiesen wird.

13. Verfahren nach einem der Ansprüche 10-12, das des weiteren einen Vervielfachungsschritt umfaßt, der vor oder  
20 in Konkurrenz mit dem Hybridisierungsschritt stattfindet.

14. Verfahren nach Anspruch 13, worin der vervielfachungsschritt PCR oder LCR umfaßt.

25 **Revendications**

**Revendications pour les Etats contractants suivants : AT, BE, CH, LI, DE, DK, FR, GB, GR, IT, NL, SE**

30 1. Composition utile dans la LCR pour amplifier l'ADN de virus du papillome humain présent dans échantillon à doser, ladite composition comprenant un ensemble de quatre sondes oligonucléotidiques, lesdits ensembles de sondes étant sélectionnés dans le groupe constitué par les ensembles d'oligonucléotides suivants :

35 **LCR5 : n° d'identification**

81 GCTGCAAACA ACTATACATG ATATAA,  
82 TTATATCATG TATAGTTGTT TGCAGC,  
83 TATTAGAATG TGTGTACTGC AAGCA,  
40 84 TGCTTGCAGT ACACACATTC TAATA ;

45 **LCR6 : n° d'identification**

85 CTTCACTGCA AGACATAGAA ATAA,  
86 TTATTTCTAT GTCTTGCAGT GAA,  
87 CCTGTGTATA TTGCAAGACA GTAT,  
50 88 TACTGTCTTG CAATATACAC AGG ;

55



## LCR7 : n° d'identification

	89	TATATTGCAA GACAGTATTG GAAC
5	90	GTTCCAATAC TGTCTTGCAA TTTA,
	91	TTACAGAGGT ATTTGAATTT GCATT,
	92	AATGCAAATT CAAATACCTC TGTA ; et

10

## LCR8 : n° d'identification

	93	GTATGGAACA ACATTAGAAC AGCA,
15	94	TGCTGTTCTA ATGTTGTTCC ATAC,
	95	ATACAACAAA CCGTTGTGTG ATTT,
	96	AAATCACACA ACGGTTTGTG GTAT.

20

2. Composition selon la revendication 1, destinée à amplifier l'ADN de virus du papillome humain de type 16 présent dans un échantillon à doser, ladite composition comprenant un ensemble de quatre sondes oligonucléotidiques, lesdits ensembles de sondes étant sélectionnés dans le groupe constitué par les ensembles d'oligonucléotides suivants :

25 LCR5 (n° d'identification 81, 82, 83 et 84) et LCR8 (n° d'identification 93, 94, 95 et 96).

3. Composition selon la revendication 1, destinée à amplifier l'ADN de virus du papillome humain de type 18 présent dans un échantillon à doser, ladite composition comprenant un ensemble de quatre sondes oligonucléotidiques, lesdits ensembles de sondes étant sélectionnés dans le groupe constitué par les ensembles d'oligonucléotides suivants :

30 LCR6 (n° d'identification 85, 86, 87 et 88) et LCR7 (n° d'identification 89, 90, 91 et 92).

4. Kit de détection de la présence d'ADN de virus du papillome humain dans un échantillon à doser, comprenant : une composition selon l'une quelconque des revendications 1 à 3, et en outre une ligase.

35

5. Kit selon la revendication 4, dans lequel ladite ligase est thermostable.

6. Composition utile dans la PCR pour amplifier l'ADN de virus du papillome humain présent dans un échantillon à doser, ladite composition comprenant :

40

une première amorce d'acide nucléique de direction sens, capable de s'hybrider au brin antisens de l'ADN de HPV, ladite amorce ayant de 10 à environ 30 nucléotides de long et une séquence sélectionnée dans le groupe constitué par les séquences suivantes :

45

## N° d'identification

	1	CAGATGTCTC	TGTGGCGGCC	TAGTG,
	6	GAATTAGTTA	GACCATTTAA	AAG,
50	7	GGGGAAACAC	CAGAATGGAT	A,
	81	GCTGCAAACA	ACTATACATG	ATATAA,
	85	CTTCACTGCA	AGACATAGAA	ATAA,
	89	TATATTGCAA	GACAGTATTG	GAAC et
55	93	GTATGGAACA	ACATTAGAAC	AGCA ; et

une deuxième amorce d'acide nucléique de direction antisens, capable de s'hybrider au brin sens de l'ADN

de HPV, ladite amorce ayant de 10 à environ 30 nucléotides de long et une séquence sélectionnée dans le groupe constitué par les séquences suivantes :

N° d'identification		
5	5	AGGTGTCAGG AAAACCAAAT TTATT,
	84	TGCTTGCACT ACACACATTC TAATA,
	88	TACTGTCTTG CAATATACAC AGG,
10	92	AATGCAAAT CAAATACCTC TGTA et
	96	AAATCACACA ACGGTTTGTT GTAT ;

pour autant que lesdites première et deuxième amorces s'hybrident à leurs brins respectifs antisens et sens à des emplacements tels que leurs extrémités 3' ne se chevauchent pas et que, dans la direction d'extension, les extrémités 5' desdites amorces soient plus espacées que les extrémités 3' desdites amorces.

7. Composition selon la revendication 6, dans laquelle lesdites première et deuxième amorces sont sélectionnées parmi les paires suivantes de séquences oligonucléotidiques (identifiées par leur numéro d'identification) :  
1 et 5, 6 et 5, 7 et 5, 81 et 84,  
85 et 88, 89 et 92, et 93 et 96.
8. Kit de détection de la présence d'ADN de virus du papillome humain dans un échantillon à doser, comprenant :  
une composition selon la revendication 6 ou 7 et en outre une polymérase.
9. Kit selon la revendication 8, dans lequel ladite polymérase est thermostable.
10. Oligonucléotide consensus pour hybridation du virus du papillome humain des types 6, 11, 16, 18, 31, 33 et 61, lequel oligonucléotide a d'environ 10 à environ 60 nucléotides de long et est sélectionné dans le groupe de séquences constitué par :

N° d'identification		
	1	CAGATGTCTC TGTGGCGGCC TAGTG,
	5	AGGTGTCAGG AAAACCAAAT TTATT,
35	6	GAATTAGTTA GACCATTTAA AAG et
	7	GGGAAACAC CAGAATGGAT A ;

et leurs compléments.

11. Oligonucléotide spécifique d'un type, destiné à déterminer la présence du virus du papillome humain de type 16, ayant une séquence sélectionnée dans le groupe constitué par :

N° d'identification		
45	81	GCTGCAAACA ACTATACATG ATATAA,
	82	TTATATCATG TATAGTTGTT TGCAGC,
	83	TATTAGAATG TGTGTACTGC AAGCA,
	84	TGCTTGCACT ACACACATTC TAATA,
	93	GTATGGAACA ACATTAGAAC AGCA,
50	94	TGCTGTTCTA ATGTTGTTCC ATAC,
	95	ATACAACAAA CCGTTGTGTG ATTT et
	96	AAATCACACA ACGGTTTGTT GTAT ;

et leurs compléments.

12. Oligonucléotide spécifique d'un type, destiné à déterminer la présence du virus du papillome humain de type 18, ayant une séquence sélectionnée dans le groupe constitué par :

## N° d'identification

	85	CTTCACTGCA	AGACATAGAA	ATAA,
5	86	TTATTTCTAT	GTCTTGCAGT	GAA,
	87	CCTGTGTATA	TTGCAAGACA	GTAT,
	88	TACTGTCTTG	CAATATACAC	AGG,
	89	TATATTGCAA	GACAGTATTG	GAAC,
	90	GTTCCAATAC	TGTCTTGCAA	TTTA,
10	91	TTACAGAGGT	ATTTGAATTT	GCATT et
	92	AATGCAAATT	CAAATACCTC	TGTAA ;

et leurs compléments.

13. Procédé de détermination de la présence d'un virus quelconque du papillome humain dans un échantillon à doser, comprenant :

- a. l'hybridation de l'ADN dans l'échantillon à doser avec au moins un oligonucléotide consensus sélectionné dans le groupe selon la revendication 10, ledit oligonucléotide étant conjugué à un composé émetteur d'un signal, capable de produire un signal détectable, et  
b. la détermination de la présence du virus du papillome humain par détection du signal émis.

14. Procédé de détermination de la présence du virus du papillome humain de type 16 dans un échantillon à doser, comprenant :

- a. l'hybridation de l'ADN dans l'échantillon à doser avec au moins un oligonucléotide sélectionné dans le groupe selon la revendication 11, ledit oligonucléotide étant conjugué à un composé émetteur d'un signal, capable de produire un signal détectable, et  
b. la détermination de la présence du virus du papillome humain par détection du signal émis.

15. Procédé de détermination de la présence du virus du papillome humain de type 18 dans un échantillon à doser, comprenant :

- a. l'hybridation de l'ADN dans l'échantillon à doser avec au moins un oligonucléotide sélectionné dans le groupe selon la revendication 12, ledit oligonucléotide étant conjugué à un composé émetteur d'un signal, capable de produire un signal détectable, et  
b. la détermination de la présence du virus du papillome humain par détection du signal émis.

16. Procédé selon une quelconque des revendications 13 à 15, comprenant en outre une étape d'amplification avant ou pendant ladite étape d'hybridation.

17. Procédé selon la revendication 16, dans lequel ladite étape d'amplification comprend la PCR ou la LCR.

#### Revendications pour l'Etat contractant suivant : ES

1. Composition utile dans la LCR pour amplifier l'ADN de virus du papillome humain présent dans échantillon à doser, ladite composition comprenant un ensemble de quatre sondes oligonucléotidiques, lesdits ensembles de sondes étant sélectionnés dans le groupe constitué par les ensembles d'oligonucléotides suivants :

## LCR5 : n° d'identification

	81	GCTGCAAACA	ACTATACATG	ATATAA,
	82	TTATATCATG	TATAGTTGTT	TGCAGC,
55	83	TATTAGAATG	TGTGTAAGTC	AAGCA,
	84	TGCTTGCAGT	ACACACATTC	TAATA ;

5  
10  
15  
20  
25  
30  
35  
40  
45  
50  
55

85 CTTCACTGCA AGACATAGAA ATAA,  
86 TTATTTCTAT GTCTTGCAGT GAA,  
87 CCTGTGTATA TTGCAAGACA GTAT,  
88 TACTGTCTTG CAATATACAC AGG ;

89 TATATTGCAA GACAGTATTG GAAC  
90 GTTCCAATAC TGTCTTGCAA TTTA,  
91 TTACAGAGGT ATTTGAATTT GCATT,  
92 AATGCAAATT CAAATACCTC TGTAAG ; et

93	GTATGGAACA	ACATTAGAAC	AGCA,
94	TGCTGTTCTA	ATGTTGTTCC	ATAC,
95	ATACAACAAA	CCGTTGTGTG	ATIT,
96	AAATCACACA	ACGGTTTGTT	GTAT.

- 30  
35  
40  
45  
50  
55

## N° d'identification

5	1	CAGATGTCTC	TGTGGCGGCC	TAGTG,
	6	GAATTAGTTA	GACCATTTAA	AAG,
	7	GGGGAAACAC	CAGAATGGAT	A,
10	81	GCTGCAAACA	ACTATACATG	ATATAA,
	85	CTTCACTGCA	AGACATAGAA	ATAA,
	89	TATATTGCAA	GACAGTATTG	GAAC et
	93	GTATGGAACA	ACATTAGAAC	AGCA ; et

15 une deuxième amorce d'acide nucléique de direction antisens, capable de s'hybrider au brin sens de l'ADN de HPV, ladite amorce ayant de 10 à environ 30 nucléotides de long et une séquence sélectionnée dans le groupe constitué par les séquences suivantes :

## N° d'identification

20	5	AGGTGTCAGG	AAAACCAAAT	TTATT,
	84	TGCTTGCACT	ACACACATTC	TAATA,
	88	TACTGTCTTG	CAATATACAC	AGG,
25	92	AATGCAAATT	CAAATACCTC	TGTAA et
	96	AAATCACACA	ACGGTTTGTT	GTAT ;

30 pour autant que lesdites première et deuxième amorces s'hybrident à leurs brins respectifs antisens et sens à des emplacements tels que leurs extrémités 3' ne se chevauchent pas et que, dans la direction d'extension, les extrémités 5' desdites amorces soient plus espacées que les extrémités 3' desdites amorces.

35 7. Composition selon la revendication 6, dans laquelle lesdites première et deuxième amorces sont sélectionnées parmi les paires suivantes de séquences oligonucléotidiques (identifiées par leur numéro d'identification) :  
1 et 5, 6 et 5, 7 et 5, 81 et 84,  
85 et 88, 89 et 92, et 93 et 96.

40 8. Kit de détection de la présence d'ADN de virus du papillome humain dans un échantillon à doser, comprenant :  
une composition selon la revendication 6 ou 7 et  
en outre une polymérase.

45 9. Kit selon la revendication 8, dans lequel ladite polymérase est thermostable.

10. Procédé de détermination de la présence d'un virus quelconque du papillome humain dans un échantillon à doser, comprenant :

45 a. l'hybridation de l'ADN dans l'échantillon à doser avec au moins un oligonucléotide consensus sélectionné dans le groupe de séquences constitué par :

## N° d'identification

50	1	CAGATGTCTC	TGTGGCGGCC	TAGTG,
	5	AGGTGTCAGG	AAAACCAAAT	TTATT,
	6	GAATTAGTTA	GACCATTTAA	AAG et
	7	GGGGAAACAC	CAGAATGGAT	A ;

55 et leurs compléments, ledit oligonucléotide étant conjugué à un composé émetteur d'un signal, capable de produire un signal détectable, et

b. la détermination de la présence du virus du papillome humain par détection du signal émis.

11. Procédé de détermination de la présence du virus du papillome humain de type 16 dans un échantillon à doser, comprenant :

a. l'hybridation de l'ADN dans l'échantillon à doser avec au moins un oligonucléotide sélectionné dans le groupe de séquences constitué par :

**N° d'identification**

81	GCTGCAAACA	ACTATACATG	ATATAA,
82	TTATATCATG	TATAGTTGTT	TGCAGC,
83	TATTAGAATG	TGTGTAAGTC	AAGCA,
84	TGCTTGCAAGT	ACACACATTC	TAATA,
93	GTATGGAACA	ACATTAGAAC	AGCA,
94	TGCTGTTCTA	ATGTTGTTCC	ATAC,
95	ATACAACAAA	CCGTTGTGTG	ATTT et
96	AAATCACACA	ACGGTTTGTT	GTAT ;

et leurs compléments,

ledit oligonucléotide étant conjugué à un composé émetteur d'un signal, capable de produire un signal détectable, et

b. la détermination de la présence du virus du papillome humain par détection du signal émis.

12. Procédé de détermination de la présence du virus du papillome humain de type 18 dans un échantillon à doser, comprenant :

a. l'hybridation de l'ADN dans l'échantillon à doser avec au moins un oligonucléotide sélectionné dans le groupe de séquences constitué par :

**N° d'identification**

85	CTTCACTGCA	AGACATAGAA	ATAA,
86	TTATTTCTAT	GTCTTGCAAGT	GAA,
87	CCTGTGTATA	TTGCAAGACA	GTAT,
88	TACTGTCTTG	CAATATACAC	AGG,
89	TATATTGCAA	GACAGTATTG	GAAC,
90	GTTCCAATAC	TGTCTTGCAA	TTTA,
91	TTACAGAGGT	ATTTGAATTT	GCATT et
92	AATGCAAATT	CAAATACCTC	TGTAA ;

et leurs compléments,

ledit oligonucléotide étant conjugué à un composé émetteur d'un signal, capable de produire un signal détectable, et

b. la détermination de la présence du virus du papillome humain par détection du signal émis.

13. Procédé selon une quelconque des revendications 10 à 12, comprenant en outre une étape d'amplification avant ou pendant ladite étape d'hybridation.

14. Procédé selon la revendication 13, dans lequel ladite étape d'amplification comprend la PCR ou la LCR.

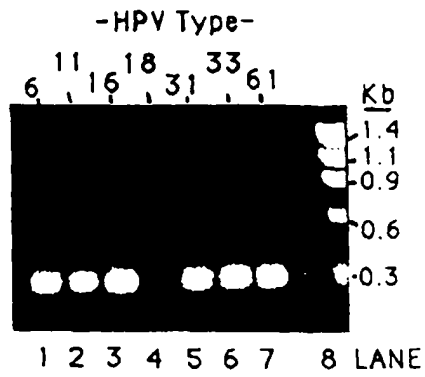


FIG. 1

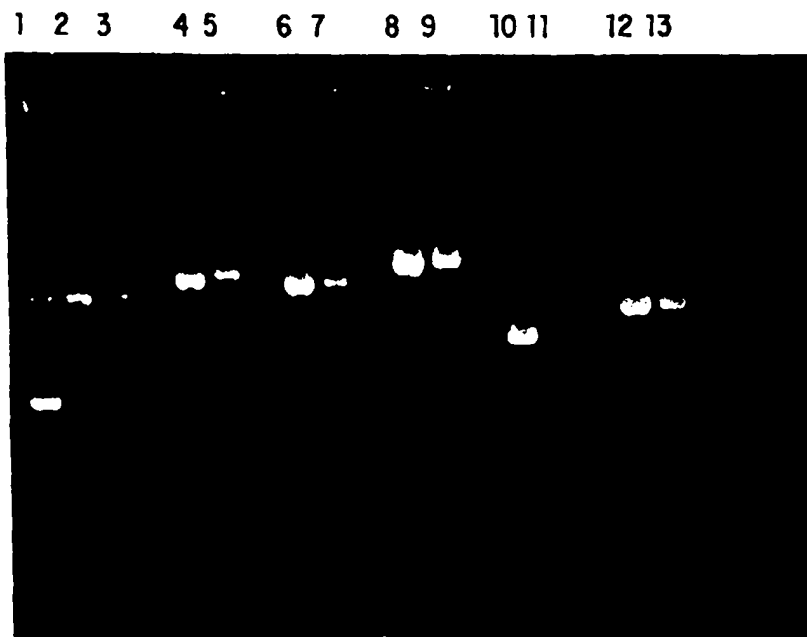


FIG. 2

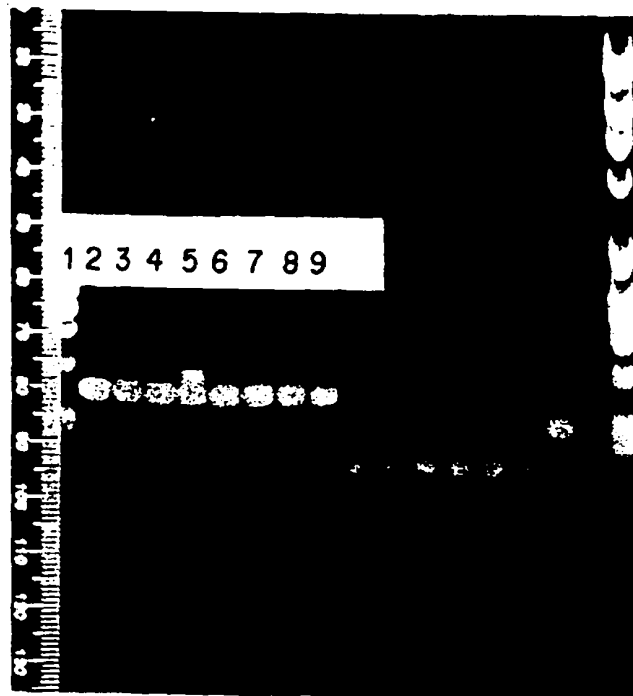


FIG. 3



FIG. 4

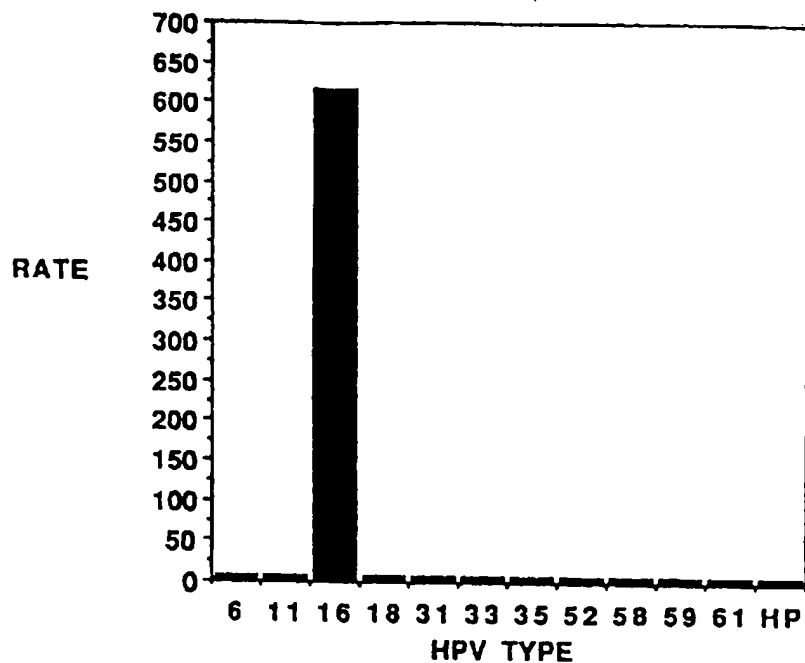


FIG. 5

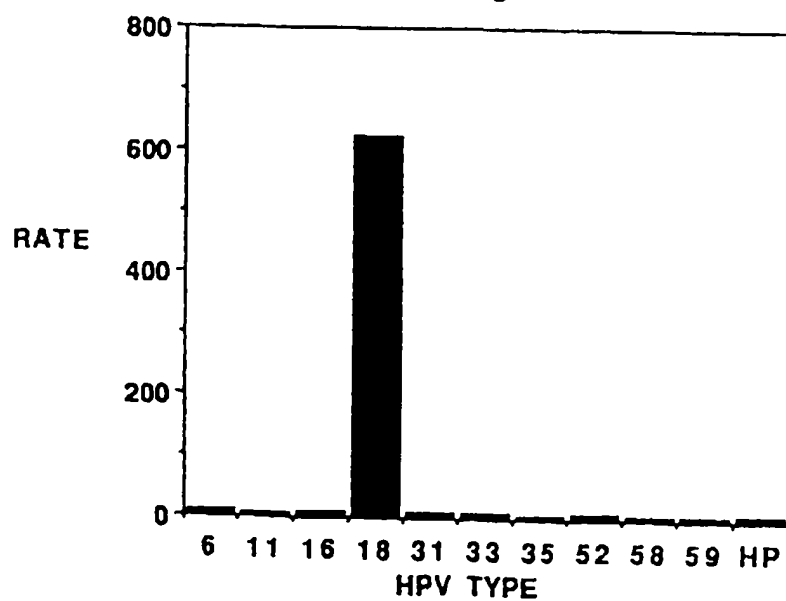


FIG. 6

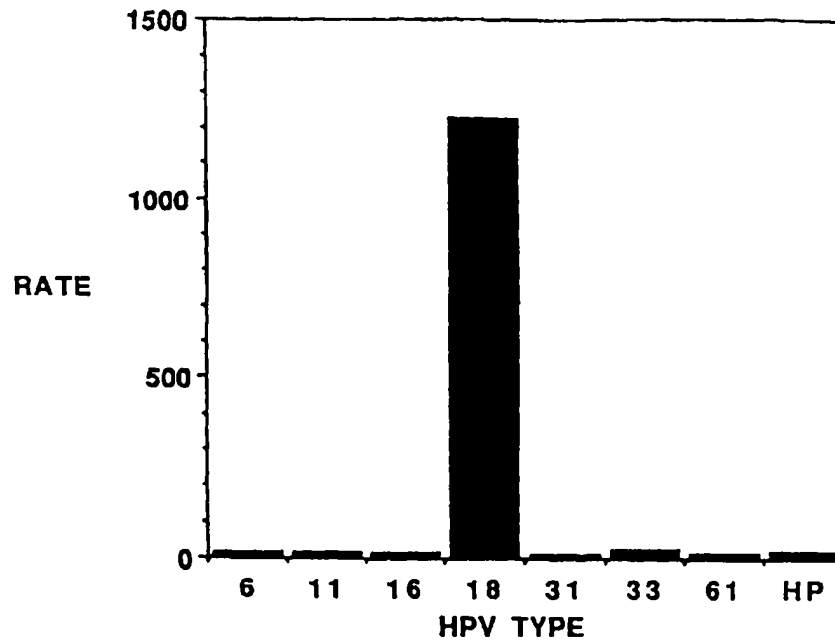


FIG. 7

